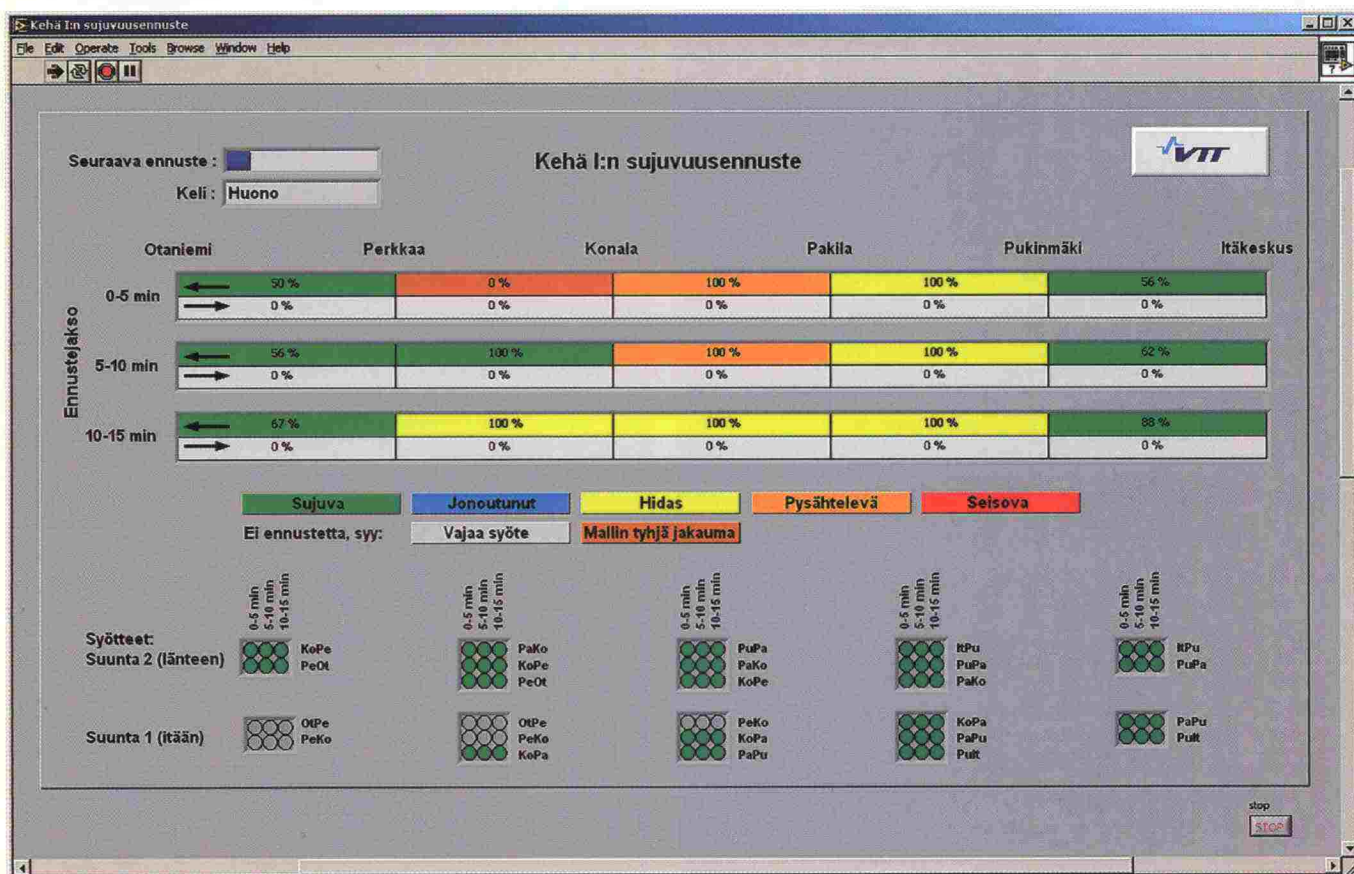


Satu Innamaa

Itseoppiva sujuvuusennuste Kehä I:lle

Tiehallinnon selvityksiä 60/2004



Satu Innamaa

Itseoppiva sujuvuusennuste Kehä I:lle

Tiehallinnon selvityksiä 60/2004

ISSN 1457-9871
ISBN 951-803-400-1
TIEH 3200910

Verkkojulkaisu pdf (www.tiehallinto.fi/julkaisut)
1459-1553
ISBN 951-803-401-X
TIEH 3200910-v

Edita Prima Oy
Helsinki2004

Julkaisua myy:
asiakaspalvelu.prima@edita.fi
Faksi 020 450 2470
Puhelin 020 450 011



Tiehallinto
Liikenteen palvelut
Opastinsilta 12 A
PL 33
00521 HELSINKI
Puhelinvaihte 0204 2211

TIIVISTELMÄ

Tutkimuksen tavoitteena oli tehdä Kehä I:lle itseoppiva lyhyen aikavälin ennustemalli, joka ennustaa tiejaksoittain seuraavan 15 minuutin kuluessa lähdössä olevien ajoneuvojen sujuvuusluokan viisiportaisen luokittelun mukaisesti. Tarkoituksena oli kehittää mallia siten, että se oppii itse kohtaamistaan liikennetilanteista ja sopeuttaa ennusteitaan niiden perusteella ilman, että kaikkea mitattua aineistoa tarvitsee tallettaa.

Tutkimuksen tavoitteet saavutettiin, sillä työn tuloksena saatiin kehitettyä itseorganisoituviin karttoihin ja klusterointiin perustuva malli, joka kykeni ennustamaan tiejaksojen sujuvuuden. Rakenteensa ansiosta malli kykeni oppimaan kohtaamistaan liikennetilanteista ilman, että kaikkea liikennetietoa piti tallettaa tietokantoihin. Tämän mahdollisti havaintojen luokittelu ja kunkin luokan vastetodennäköisyystaulukoiden päivittäminen.

Mallia tehdessä kävi selväksi, ettei pelkkä mediaani kyennyt suodattamaan kaikkia poikkeavien matka-aikahavaintojen aiheuttamia vääristymiä. Yksinkertainen menetelmä osoittautui tehokkaaksi aineistoon ajantasaisesti tehtävässä esikäsittelyssä: matka-aikamediaaniaineisto suodatettiin havaintojen lukumäärän ja sen perusteella, kuinka paljon saatu mediaani suhteellisesti poikkesi edellisestä hyväksytystä arvosta.

Online-kokeilun mukaan niiden ennusteiden osuus, jotka jäivät tekemättä siksi, että samantapaisten liikennetilanteiden klusteri oli tyhjä, pieneni odotetusti itseoppimisperiaatteen ansiosta ajan myötä. Pieneneminen oli kaikkien tiejaksojen osamalleilla suunnilleen 0,1 prosenttiyksikköä päivässä.

Malli olisi ennustanut sujuvuutta paremmin, jos sillä olisi ollut käytettävissään matka-aikojen lisäksi ajantasainen tieto liikennemääristä. Nyt liikennemäärätieto kyllä oli käytettävissä, mutta se tuli niin pitkällä viiveellä (usein jopa 20 minuuttia), että tämä viive kumosi tiedon tuomat edut.

Nyt kehitetyn mallin toimintaperiaate on sellainen, että se on sellaisenaan siirrettävissä helposti myös muualle. Kussakin kohteessa olemassa oleva liikenteen seurantajärjestelmä sanelee syötesuurevaihtoehdot. Alue, jolta yksittäisen osamallin syötteet kootaan, on riippuvainen paikasta ja sille tyypillisistä liikenteen ominaisuuksista. Syötesuureita voidaan ensin rajata karkeammin asiantuntija-arvauksen pohjalta ja lopulliset syötteet voidaan selvittää raportissa esitetyllä tavalla.

Nyckelord: Trafik information, prognos, framkomlighet

SAMMANFATTNING

Avsikten med undersökningen var att för Ring I göra en sjävlärande kortsiktig prognosmodell, som för varje vägsträcka förutspår framkomlighetsklassen för de fordon som startar inom de följande 15 minuterna enligt en modell i fem steg. Å andra sidan var avsikten att utveckla modellen så, att den lär sig själv av de trafiksituationer som den möter och anpassar sina prognoser enligt dessa, utan att allt uppmätt material behöver lagras.

Målen med undersökningen uppnåddes, ty som ett resultat lyckades man utveckla en på självorganiserande kartor och klustersampling baserad modell, som lyckades förutspå vägsträckornas framkomlighet. Tack vare sin struktur kunde modellen lära sig av de trafiksituationer den mötte, utan att man var tvungen att lagra all trafikinformation i databaserna. Detta möjliggjordes genom en klassificering av observationerna och genom att uppdatera respons sannolikhetstabellerna för varje klass.

Vid utarbetandet av modellen blev det klart, att inte enbart medianen räckte till för att filtrera bort alla snedvridningar orsakade av avvikande restidsobservationer. En enkel modell visade sig vara effektiv för förbehandlingen av materialet i realtid: restidsmedianmaterialet filtrerades på basis av antalet observationer och det erhållna medianvärdets relativa skillnad i förhållande till det föregående värdet.

Med hjälp av ett online-experiment konstaterades, att andelen sådana prognoser, som blev ogjorda på grund av att klustern med likartade trafiksituationer var tom, minskade med tiden tack vare självlärningsprincipen. Takten för alla delmodeller var ungefär 0,1 procentenheter per dag.

Modellen skulle ha förutspått framkomligheten bättre, om den förutom restiderna skulle ha haft realtidsinformation om trafikmängderna. Nu fanns trafikmängdsinformationen nog tillgänglig, men den erhöles med så lång fördröjning (ofta t.o.m. 20 minuter), att denna fördröjning upphävde nyttan.

Den nu utvecklade modellens verksamhetsprincip är sådan, att den går lätt att flytta till en annan plats också som sådan. På varje ställe dikterar trafikens uppföljningssystem alternativen för inputvariablerna. Det område, varifrån inputdata för en enskild delmodell insamlas, är beroende av platsen och dess typiska trafikegenskaper. Inputvariablerna kan först bestämmas grovt med hjälp av expertutlåtanden och de slutliga variablerna kan klarläggas på det sätt som beskrivs i rapporten.

Keywords: Traffic information, prediction, flow status

SUMMARY

The purpose of the study was to create a self-learning short-term prediction model for the flow status of the vehicles entering the road sections within next 15 minutes. The purpose was to develop a model capable of learning from the traffic situations that it observes and of adjusting the forecasts according to them without saving all the measurement data into databases.

The targets of the study were fulfilled, as the result was a model based on self-organising maps and clustering that could predict the flow status for the road sections. The structure made it possible for the model to learn from the traffic situations without a need to save all the measurements into databases. The classification and the updating of the output probability classes made it possible.

When developing the model, it became obvious that a simple median could not filter all the distortions caused by outliers in the travel time observations. A simple method turned out to be efficient in the online preprocessing of the data: travel time median data was filtered according to the number of observations and how much the current travel time median relatively differed from the previous accepted median value.

According to an online trial, the proportion of the forecasts that could not be made because the cluster of similar traffic situations was empty decreased in time due to the self-learning principle, just as expected. The decrement was for all road section approximately 0.1 percent per day.

The model could have predicted the flow status better if information on flow rates had been available in addition to travel times. Now, the flow rate information was available but with such a long delay (often up to 20 minutes) that the delay cancelled out the benefits of the additional information.

The operational principle of this model can be transferred easily to other locations as such. The traffic monitoring system of the location dictates the input candidates. The area from which the input is collected for a certain sub-model is dependent on the location and on the characteristics of traffic. Input candidates for the prediction model can be predefined according to an expert opinion but the final input variables should be selected according to the procedure explained in the report.

The project has been granted European Community financial support in the field of Trans-European Networks - Transport.

ESIPUHE

Tiehallinto sekä liikenne- ja viestintäministeriö ovat tilanneet tämän tutkimuksen itseoppivasta lyhyen aikavälin sujuvuusennustemallista osana ministeriön FITS – Liikennetelematiikan rakenteiden ja palvelujen tutkimus- ja kehittämisohjelmaa 2001–2004.

Tutkimuksesta vastasi tutkija Satu Innamaa VTT Rakennus- ja yhdyskuntatekniikasta. Lopullisen mallin ohjelmoinnin ja aineiston keräämisen teki tutkimusinsinööri Mikko Kallio VTT:stä.

Tutkimuksen etenemistä valvoi asiantuntijaryhmä, johon kuuluivat Seppo Öörni liikenne- ja viestintäministeriöstä, Sami Luoma ja Jyri Vilhunen Tiehallinnon liikenteen palvelut -yksiköstä, Mauri Pyykönen Liikennekeskuksesta, Jyri Mustonen Uudenmaan tiepiiristä sekä Risto Kulmala VTT:stä.

Tutkimuksen toteuttamiseen on saatu Euroopan unionin komission TEN-T-rahoitustukea.

Helsingissä joulukuussa 2004

Tiehallinto
Liikenteen palvelut

Sisältö

1	JOHDANTO	11
1.1	Tausta	11
1.2	Tavoitteet	12
2	MENETELMÄ	14
2.1	Tutkimuskohde	14
2.2	Aineisto	16
2.3	Itseorganisoituva kartta	17
2.4	Työkalut	18
3	TULOKSET	19
3.1	Itseorganisoituva kartta	19
3.1.1	Syötteet ja vasteet	19
3.1.2	Matka-aikamediaanien suodattaminen	20
3.1.3	Aineiston esikäsittely	20
3.1.4	Itseorganisoituvan kartan rakenne	23
3.1.5	Itseorganisoituvan kartan alustaminen	24
3.1.6	Itseorganisoituvan kartan opettaminen	24
3.2	Ennustemalli	24
3.2.1	Mallin tekeminen	24
3.2.2	Ennusteen tekeminen	26
3.2.3	Itseoppimisperiaate	27
3.3	Mallin testaus	28
3.3.1	Yleistä	28
3.3.2	Kartan koko	28
3.3.3	Luokkien erottuminen kartalla	28
3.3.4	Naapuruston koko	29
3.4	Online-kokeilun tulokset	30
3.5	Liikennemäärätiedot lisäsyötteinä	33
3.6	Kelin vaikutus	35
3.7	Viikonpäivän vaikutus	36
3.8	Puuttuva tieto	36
3.9	Lopullisen mallin ohjelmoiminen	37
4	TULOSTEN TARKASTELU JA PÄÄTELMÄT	39
	KIRJALLISUUSLÄHTEET	41
	LIITTEET	42

1 JOHDANTO

1.1 Tausta

Yksilötasolla tieosuuden tai muun liikennejärjestelmän osan liikenteen sujuvuutta arvioidaan sillä perusteella, kuinka häiriöttömästi ja odotusten mukaisesti tienkäyttäjä kykenee kulkemaan sen läpi. Tärkein yksittäinen objektiivinen liikenteen sujuvuuden mittari on matka-aika. Odotusten mukaisuudella kuvataan matkan ennustettavuutta. Liikkujien kannalta olennaista näyttää olevan kokonaismatka-ajan ohella se, kuinka paljon joudutaan poikkeamaan oletetusta tai tavoitematka-ajasta. (Luoma 1998.)

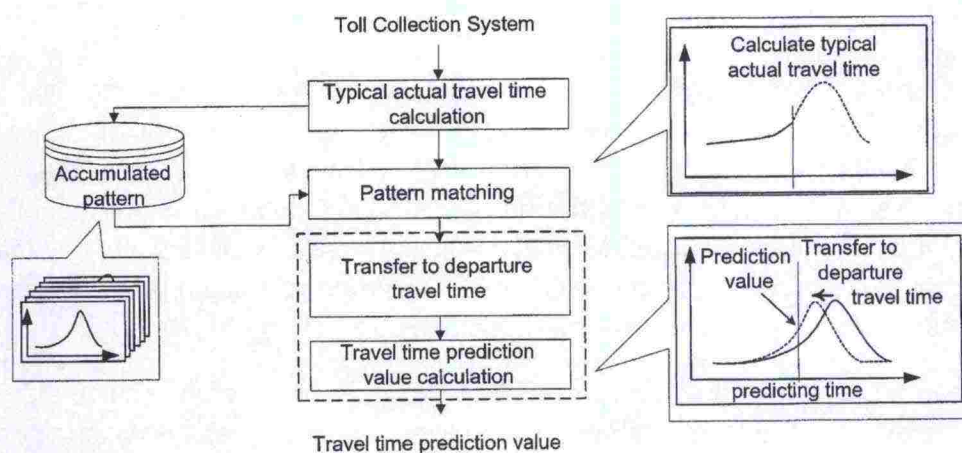
Innamaa (2004) tutki lyhyen aikavälin matka-aikaennusteen tekemistä ajantasaisesti Lahti–Heinola-välille. Tutkimuksen mukaan jopa yksinkertainen ennustemalli, joka tekee lyhyen aikavälin matka-aikaennusteita neuroverkon avulla, voi parantaa matka-aikatiedon tasoa huomattavasti verrattuna suoraan viimeisiin mittauksiin perustuviin arvioihin. Ennustemallissa oli kuitenkin ongelmana, että jos ruuhkailmiö syystä tai toisesta muuttuu, malli tarvitsee uudelleen opetuksen. Malli ei myöskään opi itsekseen kohtaamistaan satunnaisista liikenteen häiriöistä. Opetusaineisto on aina rajallinen ja mallin käsin sopeuttaminen työlästä. Tästä syystä Innamaa koki itseoppivan mallin kehittämisen tärkeäksi tulevaisuuden tutkimuskohteeksi.

Ohba ym. (2000) kehittivät hahmontunnistukseen perustuvan matka-ajan ennustemallin. Malli toimi siten, että ensin matka-aika-aineistosta poistettiin poikkeavat matka-ajat. Näitä olivat erittäin lyhyet (ruuhkan ohittavat moottoripyörät) ja pitkät (pysähtyneet ajoneuvot yms.) matka-ajat sekä hieman ja-kaumasta poikkeavat matka-ajat (kuljettajat, jotka pitävät lujaa tai hitaasti ajamisesta). Moottoripyörien matka-aikojen pois jättäminen ruuhkaliikenteestä poisti erittäin lyhyet matka-ajat. Erittäin pitkät matka-ajat eliminoitiin frekvenssianalyysin perusteella. Hieman poikkeavat matka-ajat määriteltiin klusterianalyysin perusteella ja tämän jälkeen poistettiin. Tyypillinen todellinen matka-aika määriteltiin laskemalla keskiarvo jäljelle jääneistä havainnoista.

Ohba ym. (2000) ennustivat matka-ajan seuraavan menettelyn avulla (kuva 1):

1. Laskettiin ennustepäivän matka-aikanäytteen ja kaikkien koottujen matka-aikanäytteiden väliset neliövirheet.
2. Valittiin ne matka-aikanäytteet, joiden neliövirhe oli pieni.
3. Valittiin kohdan 2 matka-aikanäytteistä jakso, joka oli ennustehetkeä edeltävältä ja seuraavalta tunnilta.
4. Valittiin lopullinen näyte, joka oli kaikkein samankaltaisin ennustehetken näytteen kanssa.

5. Järjestettiin aineisto ensimmäisen mittauspisteen ohitusajan mukaisesti.
6. Määriteltiin matka-aikaennuste kohdan 5 aineiston perusteella.



Kuva 1. Hahmontunnistukseen perustuva matka-ajan ennustemalli (Ohba ym. 2000).

Myös Otokitan ja Hashiban (1998) matka-ajan ennustemallissa sovellettiin hahmontunnistusta. Heidän mallissaan etsittiin ensin nykyolosuhteita (liikennemäärä) vastaavia tilanteita tietokannasta hahmontunnistuksen keinoin. Havainnoista etsittiin ne näytteet, jotka olivat mahdollisimman samankaltaisia kuin ennustehetken näyte ja jotka olivat mahdollisimman samalta ajanhetkeltä kuin ennustehetki. Ennustettava matka-aika määriteltiin näiden lähimpien naapureiden aineiston perusteella. Tähän aineistoon sovitettiin moniregressiomalli, joka toimi ennustemallina. Matka-aikaennuste saatiin syöttämällä malliin ennustehetken liikennemäärätiedot.

Ohban ym. (2000) ja Otokitan ja Hashiban (1998) malleissa on se ongelma, että jotta malli kehittyisi ja oppisi, sen pitäisi jatkuvasti kehittää matka-aikanäytteiden tietokantaa. Jos sinne tallennetaan kaikki havainnot, tietokanta kasvaa äkkiä erittäin suureksi, jolloin sitä on raskas käyttää. Jos taas sinne tallennetaan ainoastaan kannan näytteistä poikkeavat havainnot, kanta vääristyy.

1.2 Tavoitteet

Tämän tutkimuksen tavoitteena oli tehdä itseoppiva lyhyen aikavälin sujuvuusennustemalli Kehä I:lle, jossa on matka-ajan mittausjärjestelmä. Tarkoituksena oli tehdä malli, joka ennustaa tiejaksoittain seuraavan 15 minuutin kuluessa lähdössä olevien ajoneuvojen sujuvuusluokan viisiportaisen luokittelun mukaisesti (taulukko 1) viiden minuutin jaksoissa. Toisaalta tarkoituksena oli kehittää mallia siten, että se oppii itse kohtaamistaan liikennetilän-

teista ja sopeuttaa ennusteitaan niiden perusteella ilman, että kaikkea mitattua aineistoa tarvitsee tallettaa.

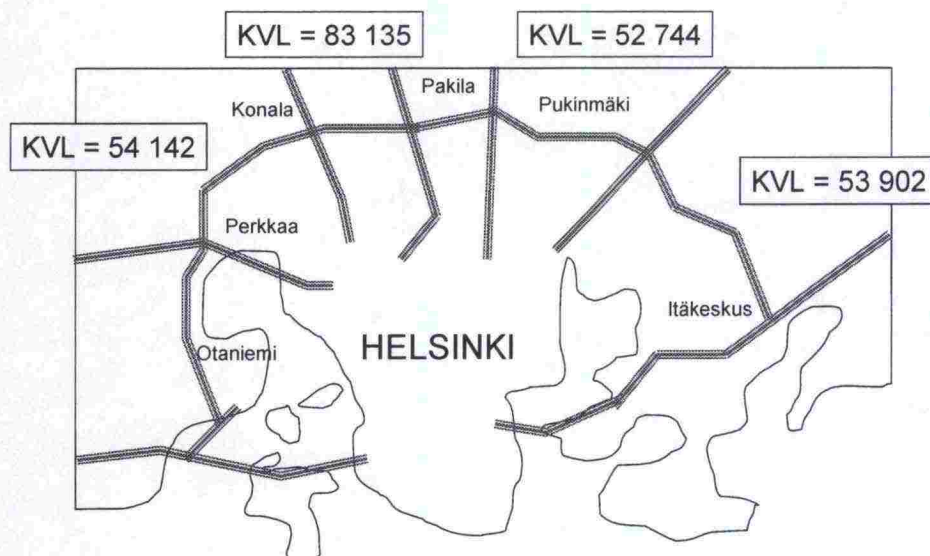
Taulukko 1. Sujuvuusluokat ja niiden määritelmät (Kiljunen & Summala 1996).

Sujuvuusluokka	Matka-nopeuden ja vapaan nopeuden osamäärä (%)
Sujuva	> 90
Jonoutunut	75–90
Hidas	25–75
Pysähtelevä	10–25
Seisova	< 10

2 MENETELMÄ

2.1 Tutkimuskohde

Tutkimuskohteena oli Kehä I (Mt 101). Tiellä on päivittäistä ruuhkaa liikennemäärän noustessa ruuhkahuippujen aikana yli 3 000 ajoneuvoon tunnissa suuntaa kohti. Vuorokauden keskimääräinen liikennemäärä on suurimmillaan yli 83 000 ajoneuvoa vuorokaudessa ja vilkkaimpien arkipäivien aikana liikennemäärä nousee yli 100 000 ajoneuvoon vuorokaudessa (kuva 2).



Kuva 2. Kehä I ja siihen liittyvät päätiet sekä Kehä I:n keskimääräiset vuorokausiliikennemäärät vuonna 2003.

Kehä I:llä on kaksi ajokaistaa suuntaan lukuun ottamatta liikenteellisesti vilkkainta jaksoa, jolla kaistoja on kolme suuntaan. Risteävien pääväylien liikenne ohjataan Kehä I:lle eritasoliittymien kautta, mutta tiellä on lisäksi myös valo-ohjauksisia tasoliittymiä katuverkon kanssa.

Kehä I:llä on kuusi ajantasaiseen keruuseen liitettyä liikenteen automaattista mittauspistettä (LAM). Nämä sijaitsevat Maxin suoralla, Konalassa, Kannelmäessä, Länsi-Pakilassa, Pukinmäessä ja Malmilla. Maxin suoran ja Kannelmäen LAM-pisteet toimivat kuitenkin epäluotettavasti vuoden 2004 alkupuoliskon aikana ja ne päätettiin jättää tutkimuksesta pois.

LAM-pisteiden lisäksi Kehä I:n liikennettä seurataan matka-aikojen automaattisen mittausjärjestelmän avulla. Matka-aikajärjestelmässä on Kehä I:llä kuusi kamerapistettä, jotka sijaitsevat Otaniemessä, Perkkaalla, Konalassa, Pakilassa, Pukinmäessä ja Itäkeskuksessa (kuva 3). Kehä I:llä ei ole yhtään ajantasaisessa keruussa mukana olevaa tiesääasemaa. Lähin tiesääasema sijaitsee Pirkkolassa valtatie 3 varrella.



Kuva 3. Matka-aikajärjestelmän kameroiden sijainnit Kehä I:llä.

Matka-aikojen seurantajärjestelmän toiminta perustuu ajoneuvojen tunnistamiseen liikennevirrasta ilman ajoneuvoihin asennettavia erillisiä tunnisteita tai lähettämiä. Ajoneuvo pyritään tunnistamaan yksilöllisesti rekisterikilven perusteella aina, kun se ohittaa järjestelmän mittauspisteen. Pisteen ohi kulkevien ajoneuvojen rekisteritunnukset ja ohitusajankohdat tallennetaan, ja peräkkäisten pisteiden tietoja verrataan keskenään. Yhdistämällä saman ajoneuvon tiedot kahdessa pisteessä saadaan laskettua ajoneuvon matkaan käyttämä aika eli ajoneuvon matka-aika. Järjestelmä perustuu Golden River Traffic Ltd:n toimittamiin laitteistoihin. (Eloranta 1999.)

Tällä hetkellä Kehä I:n sujuvuudesta annetaan tienkäyttäjille tietoja Tiehallinnon Internet-sivuilla <http://www.tiehallinto.fi/alk/frames/liikenne-frame.html>. Sujuvuustieto perustuu LAM-pisteiden viimeksi antamiin keskinopeustietoihin. Sujuvuusluokka määritellään matka-nopeuden ja vapaan nopeuden osamäärän perusteella (taulukot 1 ja 2). Luokitus perustuu Kiljusen ja Summalan (1996) tutkimuksen tuloksiin.

Taulukko 2. Tiejaksojen pituudet sekä vapaa nopeus kesä- ja talvirajoituksilla.

Tiejakso	Pituus (km)	Vapaa nopeus (km/h)	
		Kesärajoitukset	Talvirajoitukset
Otaniemi–Perkkaa	3,9	57	57
Perkkaa–Konala	3,3	67	62
Konala–Pakila	4,1	77	77
Pakila–Pukinmäki	3,1	77	77
Pukinmäki–Itäkeskus	7,4	70	70
Itäkeskus–Pukinmäki	7,4	70	70
Pukinmäki–Pakila	3,1	77	77
Pakila–Konala	4,1	77	77
Konala–Perkkaa	3,5	67	67
Perkkaa–Otaniemi	3,9	57	57

2.2 Aineisto

Tutkimus perustui pääosin vuoden 2004 tammi-elokuun aikana koottuihin matka-aikahavaintoihin. Matka-aikatieto koottiin ajantasaisesti viiden minuutin välein. Matka-aika-aineistossa oli seurantajärjestelmän aiheuttamaa keruuviivettä maksimissaan noin 240 sekuntia, sillä järjestelmän keskustietokone keräsi tiedot kameroilta 180 sekunnin välein ja tämän lisäksi aineiston prosessointiin (parien muodostaminen) kului noin 60 sekuntia.

Viimeiseltä viideltä minuutilta ennustemallille saakka saatujen havaintojen määrä oli siis usein pieni johtuen tästä viiveestä. Havaintoja saatiin kuitenkin seuraavalla tietojen hakukerralla lisää, joten aikajaksolle 5–10 minuuttia ennen ennusteen tekohetkeä saatiin usein enemmän havaintoja kuin viisi minuuttia aiemmin aikajaksolle 0–5 minuuttia ennen ennustentekohetkeä, vaikka jakso oli sama.

Matka-aikatietoa täydentämään käytettiin LAM-ilmaisimilta mitattuja liikennemääriä sekä tiesäätietoja. Tiedot haettiin viiden minuutin välein Tiehallinnon koneelta http-sovelluksen avulla. Viiden minuutin jaksoille aggregoidut pistekohtaiset liikennetiedot päivittyivät palvelimelle kymmenen minuutin välein. Aineistossa oli viiden minuutin keruuviiveen lisäksi noin kymmenen minuutin käsittely- ja tiedonsiirtoviive, ennen kuin tiedot olivat mallin käytettävissä.

Tiesääasemat tuottavat moninaista tietoa säästä ja kelistä. Aineisto luokiteltiin kelin perusteella kolmeen ryhmään: Keli luokiteltiin erittäin huonoksi, jos molemmilla kaistoilla oli lunta tai jätää ja lisäksi tienpinnan lämpötila oli alle +2 °C jommallakummalla kaistalla, jos näkyvyys oli alle 150 metriä tai jos keskituulen nopeus oli yli 16,9 m/s.

Keli luokiteltiin huonoksi, jos näkyvyys oli 150–299 metriä tai jos keskituulen nopeus oli 12–16,9 m/s. Keli luokiteltiin normaaliksi, jos molemmat kaistat olivat kuivia, kosteita, märkiä tai märkiä ja suolattuja ja sadeanturi osoittaa poutaa tai heikkoa sadetta eivätkä varoitusanturit ilmoita varoituksia. Mikäli mikään edellä esitetty ehto ei täyttnyt, keli luokiteltiin huonoksi.

Tiesääaineistossa tiedonsiirtoviivettä ei ollut käytännöllisesti katsoen lainkaan (alle 30 sekuntia). Tiedot päivittyivät palvelimelle talvikaudella noin 20 minuutin välein ja kesällä, koska huonoa keliä on vähän, noin 60 minuutin välein. Malli tulkitsee tuoreeksi tiedon, joka on korkeintaan 65 minuuttia vanha. Pitkä viive on hyväksyttävissä, koska tiesääasematieto perustuu yksittäiseen mittauspisteeseen, joka sijaitsee Kehä I:n ulkopuolella valtatiellä 3, joten suureen täsmällisyyteen kelin vaihtumisajankohdissa ei kuitenkaan päästä.

2.3 Itseorganisoituva kartta

Ennustemallin pohjana käytettiin itseorganisoituvaa karttaa (self-organising map, SOM), jota kutsutaan kehittäjänsä mukaan myös Kohosen kartaksi. Itseorganisoituva kartta on perusmuodossaan ohjaamattoman oppimisen menetelmä, jota voidaan käyttää silloin, kun luokitusta ei tunneta tai haluta käyttää. Itseorganisoituvan kartan avulla pyritään kuvaamaan aineiston luonnollista kasautumista. Lähestymistapaa kutsutaan klusterianalyysiksi, aineiston klusteroinniksi tai profiloinniksi. Ohjaamattomuus viittaa siihen, että muodostaessaan kuvaa hahmoaineistosta kartta ei perusmuodossaan käytä hyväkseen tietoa hahmojen luokituksesta. (Luku 2.3 perustuu viitteisiin Kohonen 2001, Vesanto ym. 2000 ja Raitio 1998.)

Itseorganisoituva kartta muodostuu hilaan järjestetyistä karttayksiköistä. Karttayksiköiden välistä etäisyyttä voidaan mitata niiden hilakoordinaatistossa ilmoitettujen paikkavektorien etäisyyttä käyttäen. Jokaisella karttayksiköllä on siis paikkavektori, jonka ulottuvuus on sama kuin kartan syöteenä käytettävällä hahmovektorilla. Paikkavektorit liittävät jokaiseen karttayksikköön vastinpisteen hahmoavaruudessa ja vastaavasti jokaisen hahmovektoriin karttayksikön, jonka paikkavektori on hahmovektoria lähinnä.

Itseorganisoituvilla kartoilla on kaksi erityistä ominaisuutta. Paikkavektoreiden jakauma pyrkii seuraamaan opetuksessa käytetyn hahmoaineiston jakaumaa, ja hahmontunnistuksessa lähekkäin olevat paikkavektorit pyrkivät sijoittumaan hilakartalla lähekkäin oleviin karttayksiköihin. Ensimmäisen ominaisuuden perusteella karttaa voidaan käyttää yleistämään aineistoa pienellä määrällä kasoja, joita esitetään kartan paikkavektoreilla. Kasa muodostuu hahmovektoreista, jotka kuuluvat paikkavektorin ympäristöön. Toisen ominaisuuden perusteella samankaltaiset kasat ja hahmovektorit kuvautuvat lähekkäisiin karttayksiköihin.

Kyseessä on siis eräänlainen hahmoavaruuden projektio karttahilalle. Tämä ominaisuus helpottaa kasojen tulkintaa erityisesti, kun karttahila on kaksiulotteinen taso, joka voidaan visualisoida. Tyypillisesti ominaisuuksiltaan erilaiset kasat kuvautuvat eri osiin karttahilaa ja samankaltaiset kasat muodostavat yhtenäisiä alueita. Tutkimalla, millaisia hahmoja kartan eri osiin kuvautuu, saadaan käsitys siitä, mitkä ovat aineistossa hahmoja ja kasoja erottavat tekijät.

Itseorganisoituva kartta voidaan opettaa kahdella erilaisella tavalla: joko järjestystä noudattavalla tai joukkoihin perustuvalla opetusalgoritmillä. Järjestystä noudattavassa opetusalgoritmissa (Sequential training algorithm) kartta opetetaan iteratiivisesti. Jokaisella opetuskerralla syöteaineistosta valitaan satunnaisesti yksi näytevektori ja määritellään tämän vektorin etäisyys kartan paikkavektoreihin. Lähinnä olevaa paikkavektoria siirretään syötevektorin suuntaan syöteavaruudessa. Opettaminen tehdään kahdessa vaiheessa:

ensin haetaan karkeasti kartan hahmo ja tämän jälkeen paikkavektorien arvot hienosäädetään paikalleen.

Myös joukkoihin perustuva opetusalgoritmi (Batch training algorithm) on iteratiivinen, mutta sen sijaan, että käytettäisiin kerrallaan yksittäistä syötevektoria, neuroverkolle esitellään koko aineisto, ennen kuin tehdään paikkavektorien päivitykset. Jokaisella opetuskerralla aineisto jaetaan osiin siten, että kukin syötevektori kuuluu sen karttayksikön aineistoon, joka on sitä lähinnä.

Periaatteessa itseorganisoituva kartta muodostetaan valvomattoman prosessin avulla. On kuitenkin havaittu, että kartan kyky toimia luokittimena paranee, jos luokkatietoa voidaan hyödyntää opetusprosessissa. Tällaisessa ns. valvotussa opettamisessa hahmovektoreihin lisätään nollista ja ykkösistä koostuva osio, jossa ykkönen osoittaa luokan, johon vektori kuuluu. Näin tiettyyn luokkaan kuuluvat havainnot päätyvät todennäköisemmin samaan osaan karttaa kuin ilman luokkatietoa. Opetusprosessin jälkeen luokkatieto poistetaan vektoreista ja karttaa voidaan käyttää kuten tavallisesti.

2.4 Työkalut

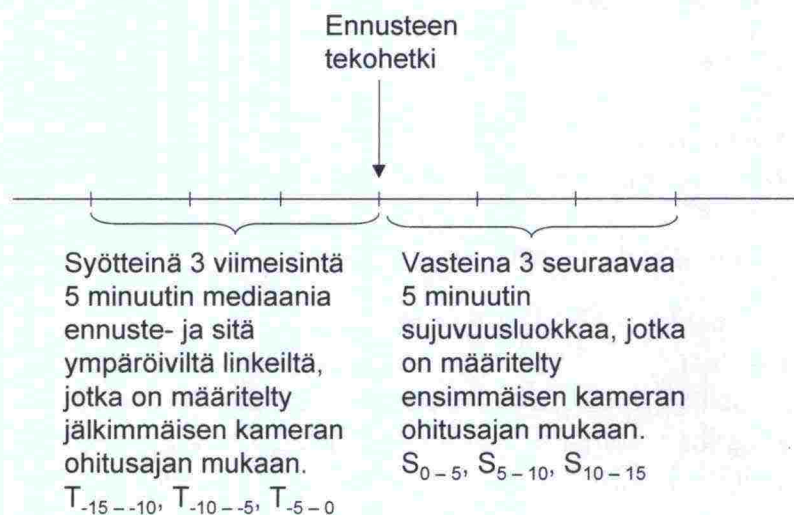
Aineiston keruussa, kokoamisessa ja esikäsittelyssä käytettiin LabView-ohjelmistoa. Ennustemallin itseorganisoituvat kartat opetettiin Matlab-ohjelmiston Somtoolbox-työkalua (Vesanto ym. 2000) hyödyntäen. Lopullinen ennusteohjelmisto käyttöliittymineen ohjelmoitiin LabView-ohjelmistolla.

3 TULOKSET

3.1 Itseorganisoituva kartta

3.1.1 Syötteet ja vasteet

Ensimmäisessä vaiheessa matka-aika-aineistosta koottiin opetusaineisto ja testiaineisto, joissa syötehavaintoihin liitettiin tieto mitatusta vasteesta. Syötteinä käytettiin ennustelinkiltä sekä sitä edeltävältä ja seuraavalta linkiltä mitattua kolmea viimeistä viiden minuutin matka-aikamediaania. Vasteena malli kertoi sujuvuusluokkien todennäköisyydet seuraavan 15 minuutin aikana ennustelinkille lähdössä oleville ajoneuvoille (kuva 4).



Kuva 4. Mallin syötteet ja vasteet aikajanalla.

Itseorganisoituva kartta muotoutuu opetuksen kuluessa esittämään tyypillisiä havaintoja. Käytännössä tärkeiden tapausten (tässä: ruuhkien) osuus opetusjoukosta saattaa olla pieni. Tällöin voi käydä niin, etteivät ne saa vallattua kartalta lainkaan omaa alaa. Tästä syystä onkin tärkeää, että näitä havaintoja toistetaan opetusaineistossa satunnaisessa järjestyksessä riittävän monta kertaa opetuksen aikana tai että niitä painotetaan muuten riittävän voimakkaasti.

Opetusjoukko koottiin siten, että kaikista sujuvuusluokista oli yhtä monta havaintoa. Tämä toteutettiin monistamalla alempien sujuvuusluokkien (jonoutunut, hidas, pysähtelevä, seisova) havaintoja riittävän monta kertaa satunnaisessa järjestyksessä, jotta havaintoja kertyi ensi vaiheessa yhtä monta kuin sujuvan liikenteen luokassa, joka oli alun perin suurin. Myöhemmin, kun tutkimuksen lopullisia malleja tehtiin, havaintoja koottiin joka luokasta satunnaisesti keräten 4 000 kappaletta.

Myöhäisemmässä vaiheessa ennusteen tekemistä kokeiltiin myös sellaisella aineistolla, jossa oli matka-aikatietojen lisäksi liikennemäärätiedot vastaavina aikasarjoina ja vastaavalta alueelta kuin matka-ajat.

3.1.2 Matka-aikamediaanien suodattaminen

Matka-aikamediaanit määriteltiin automaattisen matka-ajan mittausjärjestelmän tuottamasta raaka-aineistosta. Aineisto sisälsi mittausvirheistä tai poikkeavaa reittiä ajaneista ajoneuvoista aiheutuneita poikkeavia havaintoja. Poikkeavien havaintojen osuus oli yleensä pieni, joten ne eivät juuri vaikuttaneet mediaaniin silloin, kun se laskettiin suuresta määrästä havaintoja. Havaintomäärät eivät kuitenkaan aina olleet riittävän suuria vaimentamaan poikkeavien havaintojen vaikutusta ja mediaaniarvot päätettiin suodattaa ennen mallille syöttämistä.

Poikkeavien mediaanien suodattamisessa kokeiltiin kahta eri menetelmää: polynomin sovitusta ja peräkkäisten havaintojen suhteellisen eron maksimia. Ensin kokeiltiin suodattamista siten, että mediaanihavaintoihin sovitettiin polynomi, josta etäällä olevat havainnot voitaisiin tulkita poikkeaviksi. Menetelmää ei kuitenkaan voitu soveltaa ajantasaisesti silloin, kun havaintoja tulee vähän, joten se päätettiin hylätä.

Toisena vaihtoehtona kokeiltiin yksinkertaista menetelmää. Tässä etsittiin aluksi ne mediaanit, jotka perustuivat pieneen havaintomäärään. Sitten niistä suodatettiin pois sellaiset, jotka poikkesivat edellisestä hyväksytystä arvosta liikaa. Sopivia määritelmiä "pienelle otoskoolle" ja "liikaa poikkeamiselle" etsittiin visuaalisin keinoin kokeilemalla erilaisia raja-arvoja ja tarkkailemalla havaintoja, jotka eri tilanteissa karsiutuisivat aineistosta. Tarkoituksena oli löytää sellaiset raja-arvot, jotka löytäisivät lähes kaikki ihmissilmin poikkeaviksi tulkittavat havainnot, mutta eivät poistaisi esimerkiksi ruuhkien alkua.

Menetelmä tuntui toimivan hyvin. Paras tulos saatiin siten, että niistä mediaaneista, jotka perustuivat korkeintaan kahteen havaintoon, hyväksyttiin ainoastaan ne, jotka poikkesivat edellisestä hyväksytystä arvosta korkeintaan 50 prosenttia.

3.1.3 Aineiston esikäsittely

Aineisto esikäsiteltiin itseorganisoituvan kartan tekemistä varten, jottei mikään syötemuuttuja päässyt hallitsemaan ylivoimaisesti muihin muuttujiin verrattuna, eli jottei tässä tapauksessa matka-aikojen tärkeys painottunut tiejakson (matka-ajan) pituuden mukaisesti.

Aineiston esikäsittely voidaan toteuttaa usealla eri tavalla. Nyt kokeiltiin kolmea eri vaihtoehtoa: (1) muuttujat normeerattiin siten, että keskiarvoksi tuli nolla ja keskihajonnaksi yksi, (2) muuttujat skaalattiin nollan ja yhden välille ja (3) niille tehtiin logaritminen muunnos. Matemaattisesti nämä tehtiin seuraavasti (Vesanto ym. 2000):

$$x' = \frac{x - \bar{x}}{\sigma_x} \quad (1)$$

missä x' oli muuttujan normeerattu ja x sen alkuperäinen arvo, \bar{x} oli muuttujan x keskiarvo ja σ_x sen keskihajonta.

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (2)$$

missä $\min(x)$ oli muuttujan x minimiarvo ja $\max(x)$ sen maksimiarvo.

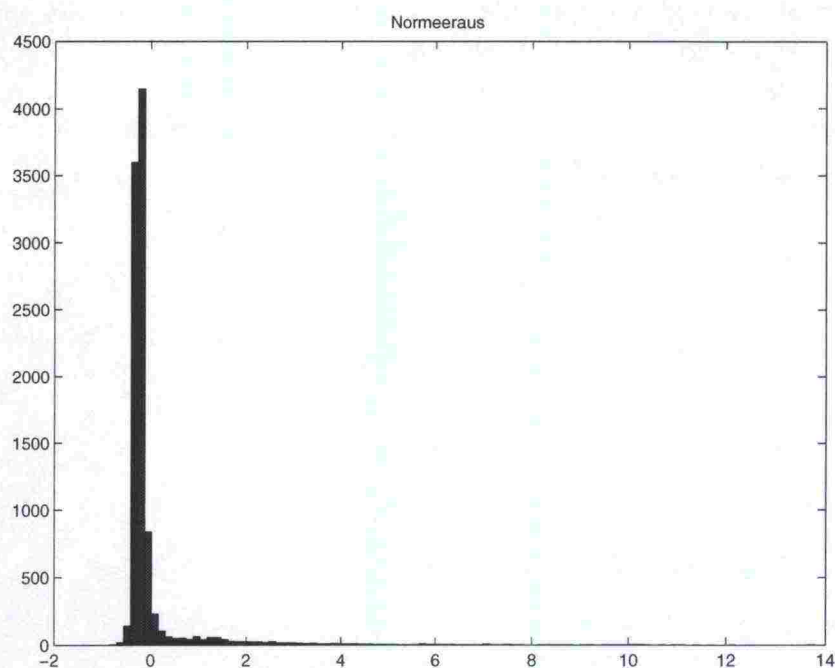
$$x' = \ln(x - (\min(x) - 1)) \quad (3)$$

missä \ln oli luonnollinen logaritmi.

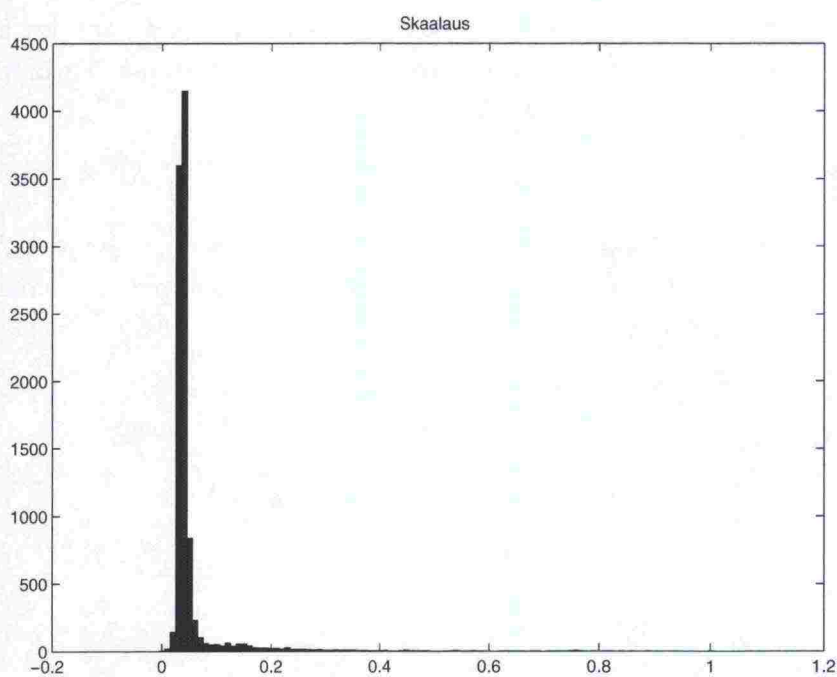
Viimeistä menetelmää käytettäessä piti varmistua siitä, ettei muunnosta tehdä pienemmille arvoille kuin $(\min(x)-1)$ silloin, kun esikäsittely tehdään ennustemallia käytettäessä uusille havainnoille, koska muutoin tuloksena on kompleksiluku. Käytännössä ongelma ratkaistiin siten, että $\min(x)$:n arvoksi asetettiin 1, jota pienemmät matka-ajat eivät ole mahdollisia. Tällöin kaava supistui muotoon

$$x' = \ln(x)$$

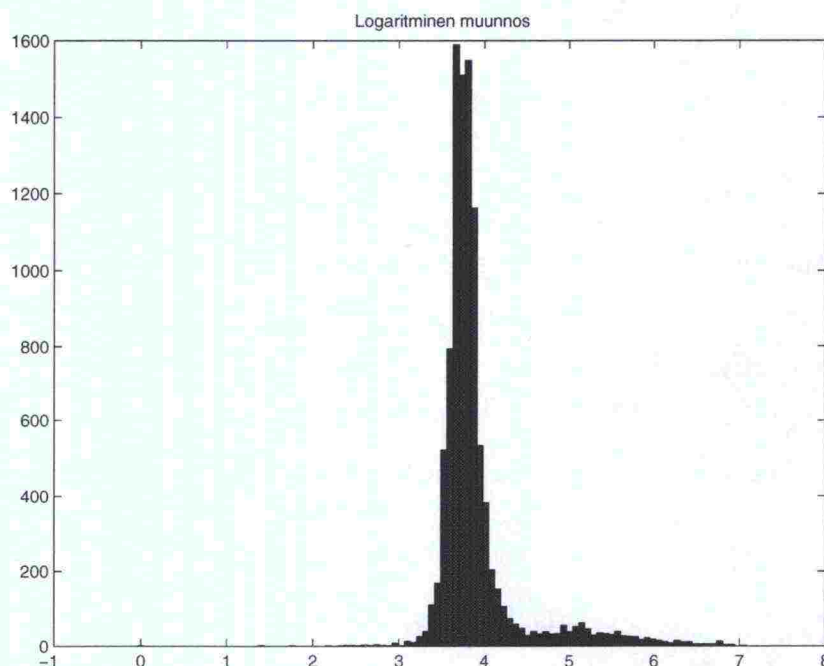
Suodatettu matka-aika-aineisto esikäsiteltiin siis kaikilla kolmella menetelmällä. Logaritminen muunnos erotteli havainnot paremmin kuin normeeraus tai skaalaus (kuvat 5–7). Tästä syystä se valittiin aineiston esikäsittelymenetelmäksi.



Kuva 5. Histogrammi normeeratusta matka-aika-aineistosta.



Kuva 6. Histogrammi skaalatusta matka-aika-aineistosta.



Kuva 7. Histogrammi logaritmisella muunnoksella käsitellystä matka-aika-aineistosta.

3.1.4 Itseorganisoituvan kartan rakenne

Itseorganisoituvan neuroverkon karttayksiköiden määrä (*Munits*) määritettiin seuraavalla heuristisella kaavalla (Vesanto ym. 2000):

$$Munits = 5 \cdot dlen^{0,54321}$$

jossa *dlen* on opetusaineiston koko. 'Iso' kartta määriteltiin neljä kertaa peruskartan kokoiseksi ja 'pienen' kartan koko neljäsosaksi peruskartasta.

Karttayksiköiden määrän jälkeen määriteltiin kartan sivujen pituudet. Tämä tehtiin käytännössä siten, että laskettiin opetusaineiston kaksi suurinta ominaisvektoria ja kartan sivujen pituuksien suhde määräytyi näiden vektorien suhteen perusteella. Sivujen pituudet määritettiin siten, että niiden tulo olisi mahdollisimman lähellä edellä haluttua karttayksiköiden määrää. (Vesanto ym. 2000.)

Kartan hilarakenne voi olla esimerkiksi kuusikulmainen, suorakulmainen tai satunnainen. Nyt rakenteeksi valittiin kuusikulmainen muoto.

3.1.5 Itseorganisoituvan kartan alustaminen

Kartan alustus voidaan tehdä satunnaisesti, sillä on osoitettu, että alunperin ei-järjestetyt vektorit järjestyvät ajan mittaan, tavallisesti muutaman sadan opetuskierroksen aikana. Tämä ei kuitenkaan tarkoita sitä, että satunnaisluvuilla alustaminen olisi paras tai nopein menettelytapa tai että sitä kannattaisi soveltaa käytännössä. (Kohonen 2001.)

Kartan alustaminen voidaan tehdä satunnaislukujen lisäksi myös valitsemalla opetusaineistosta satunnaisesti joukko havaintoja ja käyttämällä näitä kartan alkuarvoina. (Kohonen 2001.)

Kolmas menetelmä on lineaarinen alustus. Menetelmässä aineistosta määritellään ensin ominaisarvot ja -vektorit. Kartta alustetaan tämän jälkeen opetusaineiston kovarianssimatriisin suurimpien ominaisvektorien mukaisesti. Kokemus (Kohonen 2001) on osoittanut menetelmän toimivaksi ja tästä syystä kartta alustettiin sen mukaisesti. Jos tämä ei ollut jossain tapauksessa mahdollista (ominaisvektoreja ei voitu määrittää), alustus tehtiin satunnaisesti.

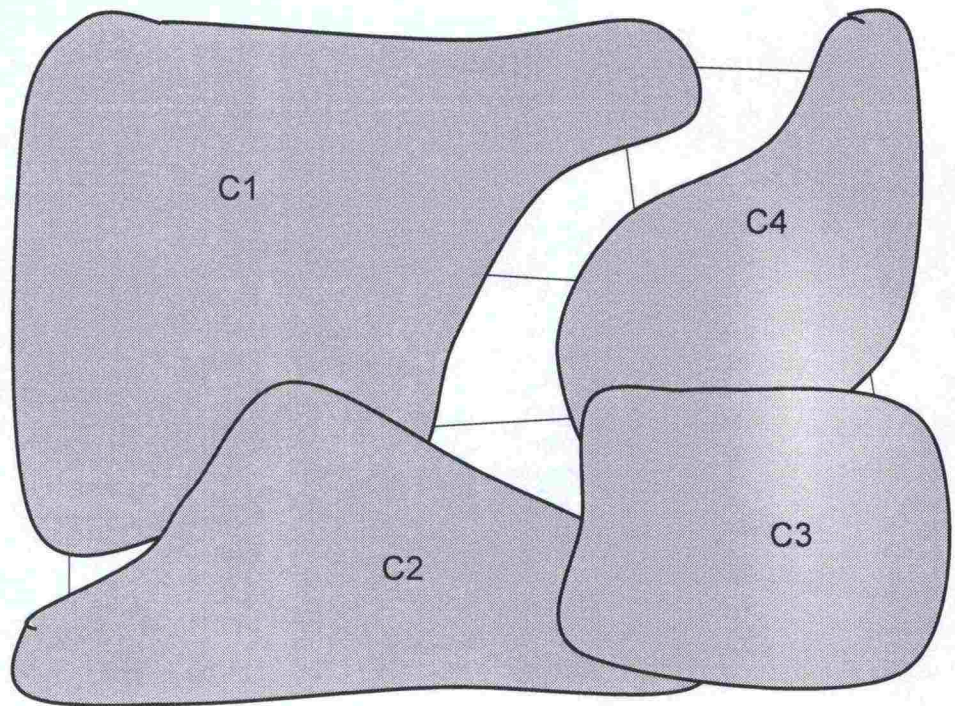
3.1.6 Itseorganisoituvan kartan opettaminen

Itseorganisoituvat kartat opetettiin valvotun opetuksen periaattein hyödyntämällä olemassa olevaa opetusaineiston luokkatietoa. Opetuksessa sovellettiin joukkoihin perustuvaa opetusalgoritmia (Batch training algorithm).

3.2 Ennustemalli

3.2.1 Mallin tekeminen

Kullekin linkkivälille (5 kpl/suunta) tehtiin itseorganisoituvat kartat kullekin kolmelle ennustejaksolle (0–5, 5–10 ja 10–15 minuutin kuluttua linkin aloittavien ajoneuvojen sujuvuusluokka). Tavoitteena oli, että eri luokkiin kuuluvat havainnot sijoittuisivat mahdollisimman hyvin eri osiin karttaa (eri klustereihin, kuva 8).

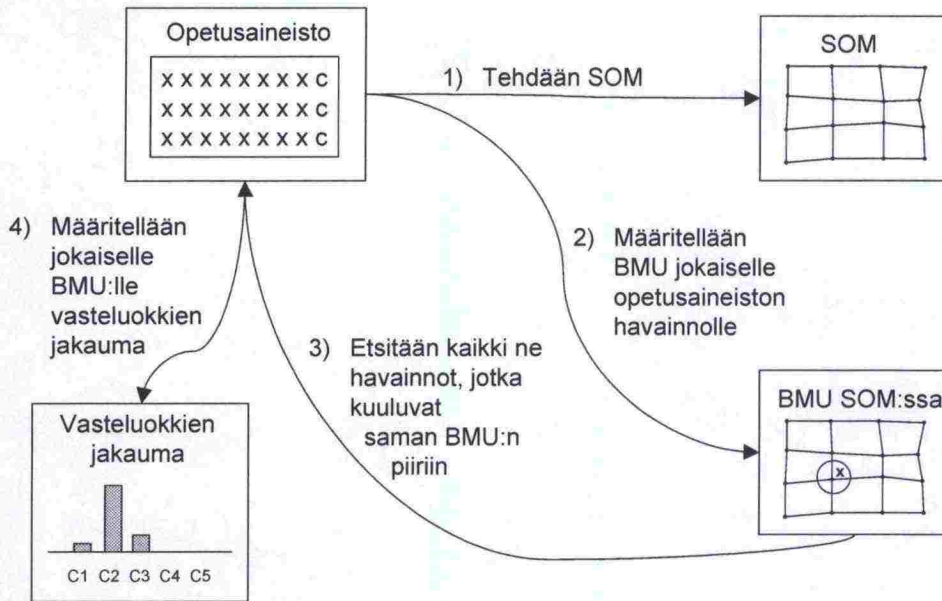


Kuva 8. Eri luokkiin kuuluvien havaintojen sijoittuminen itseorganisoituvalle kartalle. Ihannetapauksessa luokat eivät menisi lainkaan päällekkäin.

Kartan tekemisen jälkeen tutkittiin, mikä neuroni oli lähinnä mitäkin opetusaineiston havaintoa (kuva 9). BMU (best matching unit) oli se neuroni, jonka paikkavektorin (euklidinen) etäisyys kyseiseen havaintoon oli pienin. Euklidinen etäisyys vektorin $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_n]$ ja $\mathbf{y} = [y_1 \ y_2 \ \dots \ y_n]$ välillä määritellään

$$Euc(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

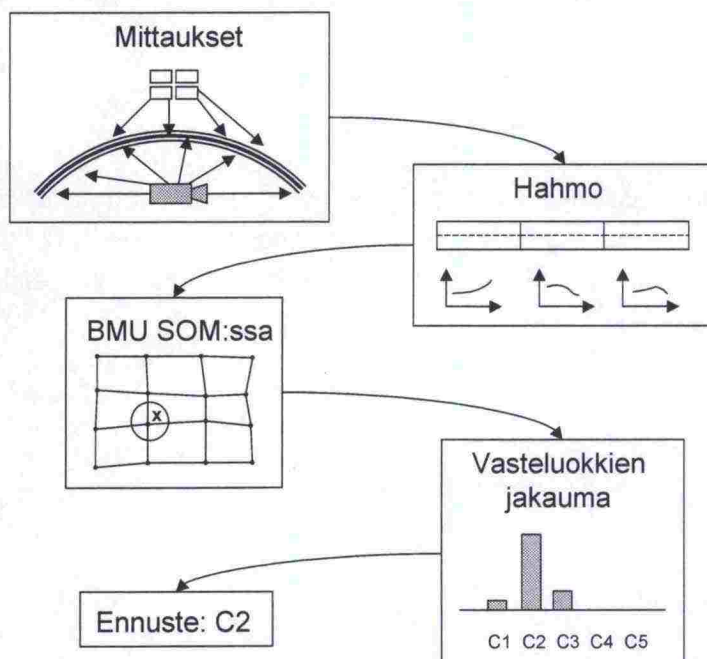
Tämän jälkeen laskettiin, mikä on kunkin vasteen todennäköisyys kussakin klusterissa. Näin saatiin eri tulemien todennäköisyyksien taulukko. Kartan onnistuneisuutta arvioitiin tarkastelemalla oikein ennustettujen sujuvuusluokien osuuksia.



Kuva 9. Ennustemallin tekoperiaate.

3.2.2 Ennusteen tekeminen

Ennuste tehtiin muuntamalla syöte logaritmisella muunnoksella ja etsimällä sille BMU, minkä jälkeen eri sujuvuusluokkien todennäköisyydet voitiin katsoa kyseisen BMU-karttayksikön todennäköisyystaulukosta (vasteluokkien jakauma, kuva 10). Ennusteeksi valittiin luokista todennäköisin.

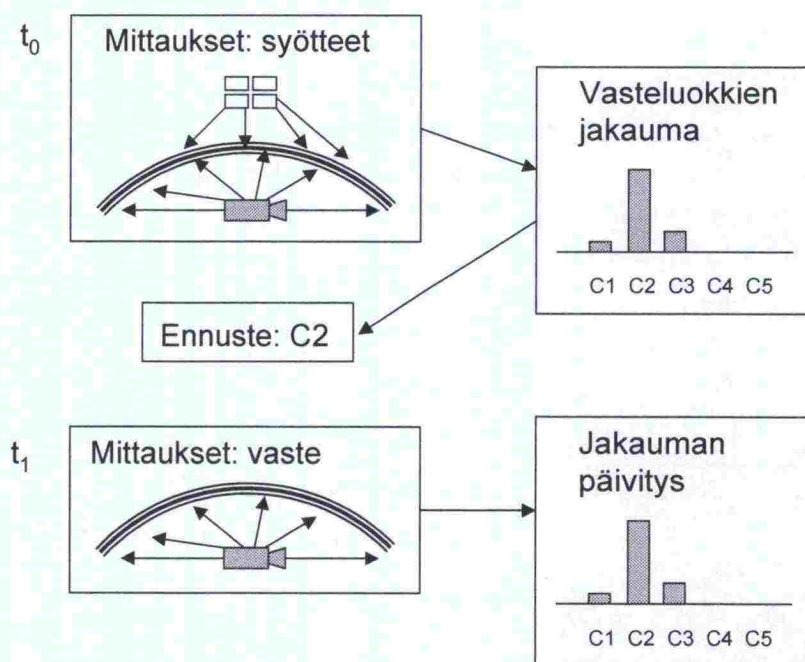


Kuva 10. Ennusteen tekoperiaate.

3.2.3 Itseoppimisperiaate

Monen mallin itseoppimisperiaate perustuu siihen, että se tallentaa tiedon kaikista kohtaamistaan liikennetilanteista ja muokkaa ennusteitaan niiden mukaisesti. Jos kaikki syötteet ja vasteet tallennetaan, vie syntyvä tietokanta ajan mittaan valtavasti levytilaa ja mallin käyttö täten hankaloituu. Jos taas tallennetaan ainoastaan ns. mielenkiintoiset tai poikkeavat tilanteet, mallin kuva liikenteestä vääristyy eikä se enää toimi oikein ns. tavallisissa tilanteissa.

Jotta voitaisiin välttää sekä levytila- että liikennetilannekuvan vääristymisongelmat, syötteen ja vasteen sijaan päätettiin tallentaa ainoastaan tieto siitä, kuinka moni havainto osui kuhunkin sujuvuusluokkaan kunkin karttayksikön alueella. Näin todennäköisyystietoja päivitettiin jokaisen saadun ns. oikean vastauksen perusteella (kuva 11).



Kuva 11. Vasteluokkien päivitysperiaate. Ennuste tehdään hetkellä t_0 ja jakaumaa päivitetään, kun todellinen vaste on saatu mitattua (hetki t_1).

3.3 Mallin testaus

3.3.1 Yleistä

Mallin toimintaperiaatteita testattiin kahdella pienellä aineistolla. Opetusaineisto oli koottu 27.1.–25.3.2004 ja testiaineisto 25.3.–14.4.2004. Monia periaatteita testattiin ensin tiejaksolla Pakila–Konala, joka on eräs Kehä I:n pahiten ruuhkautuvista osuuksista.

Aineistoa koottiin samalla, kun tehtiin mallin kehitystyötä. Tästä syystä mallin toimintaa testattiin yllä mainitulla pienellä aineistolla. Tämän alustavan, pienellä aineistolla kootun mallin käyttäytymistä selvitettiin ajantasaisessa toiminnassa. Kun mallin toimintaperiaatteet oli saatu määriteltä, varsinainen ennustemalli tehtiin kaikella aineistolla, jota siihen mennessä oli kerätty.

3.3.2 Kartan koko

Itseorganisoituvan kartan koon vaikutusta selvitettiin kolmen eri kokoisen kartan avulla. Luokka 2 (jonoutunut) vaikutti sekoittuvan luokkiin 1 (sujuva) ja 3 (hidas) sitä enemmän, mitä pienempi kartta oli kyseessä (taulukko 3). Tästä syystä ennustemalleissa päätettiin käyttää ns. suurta karttaa.

Taulukko 3. Kartan koon vaikutus. Suuren SOM:n koko oli 110 x 35, keskikokoisen 57 x 17 ja pienen 27 x 9. Ennustelinkkinä oli Pakila–Konala, ennustejakso oli 0–5 minuuttia ja aineistona oli opetusaineisto.

Ennustettu luokka	Suuri SOM					Keskikokoinen SOM					Pieni SOM				
	Mitattu luokka														
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1	8939	63	15	0	0	8916	116	39	0	0	8880	163	74	0	0
2	3	206	0	0	0	9	121	3	0	0	11	42	8	0	0
3	2	1	680	0	0	19	33	653	0	0	53	65	607	7	0
4	0	0	0	48	0	0	0	0	48	0	0	0	6	41	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

3.3.3 Luokkien erottuminen kartalla

Ihanteellisessa tapauksessa kaikki ennusteet vastaavat mitattuja luokkia (kaikki havainnot ovat soluissa $X = X$). Tässä opetusaineistossa paras luokkien erottelu oli tiejaksolla Pakila–Pukinmäki ja huonoin tiejaksolla Konala–Perkkaa (taulukot 4 ja 5).

Taulukko 4. Opetusaineiston luokat Pakila–Pukinmäki-tiejaksolla.

Ennustettu luokka	Ennustejakso														
	0–5 min					5–10 min					10–15 min				
	Mitattu luokka														
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1	5882	0	0	0	0	5886	2	0	0	0	5886	6	1	0	0
2	0	18	0	0	0	0	12	0	0	0	0	8	0	0	0
3	0	0	10	0	0	0	0	9	0	0	0	0	7	0	0
4	0	0	0	2	0	0	0	0	3	0	0	0	0	4	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Taulukko 5. Opetusaineiston luokat Konala–Perkkaa-tiejaksolla.

Ennustettu luokka	Ennustejakso														
	0–5 min					5–10 min					10–15 min				
	Mitattu luokka														
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1	7058	253	15	0	0	7059	246	36	0	1	7074	262	88	1	2
2	19	393	3	0	0	29	379	13	0	0	21	354	12	0	0
3	6	11	1286	0	0	10	16	1256	0	0	15	17	1198	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1

3.3.4 Naapuruston koko

Naapuruston koon (BMU:iden lukumäärä) vaikutus ennusteiden tarkkuuteen selvitettiin käyttämällä ns. testiaineistoa. Niiden tilanteiden määrä, jolloin ennustetta ei voitu antaa (klusterissa ei ollut havaintoja, joiden pohjalta tehdä ennustetta) väheni nopeasti, kun naapuruston koko kasvoi (taulukko 6), mikä on luonnollista. Tyhjien klusterien ongelman pitäisi kuitenkin väistyä joka tapauksessa ajan mittaan itseoppimisperiaatteen mukaisesti. Niinpä olennaisempi tulos olikin se, että luokat alkoivat sekoittua, kun naapuruston koko alkoi kasvaa. Tästä syystä naapuruston koko päätettiin pitää pienenä (1 BMU).

Taulukko 6. Naapuruston koon (BMU:iden lukumäärä) vaikutus ennusteeseen. Ennustelinkkinä oli Pakila–Konala, ennustejakso oli 0–5 minuuttia ja aineistona oli testiaineisto.

Ennustettu luokka	1 BMU					3 BMU:ta					5 BMU:ta				
	Mitattu luokka														
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	1	1114	2	0	0	0	3244	28	15	0	0	3257	45	40	0
2	0	19	45	0	0	2	23	48	0	0	10	8	27	0	0
3	0	5	14	0	0	4	7	75	0	0	4	9	93	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ei	2157	37	106	0	0	21	5	27	0	0	0	1	0	0	0

Taulukossa 6 eri BMU:iden vastejakaumat saivat yhtä suuren painoarvon riippumatta siitä, kuinka etäällä ne olivat havainnosta. Ennusteen tekemistä kokeiltiin painottamalla jakaumia etäisyyden käänteisarvolla. Tulokset eivät kuitenkaan parantuneet (taulukko 7).

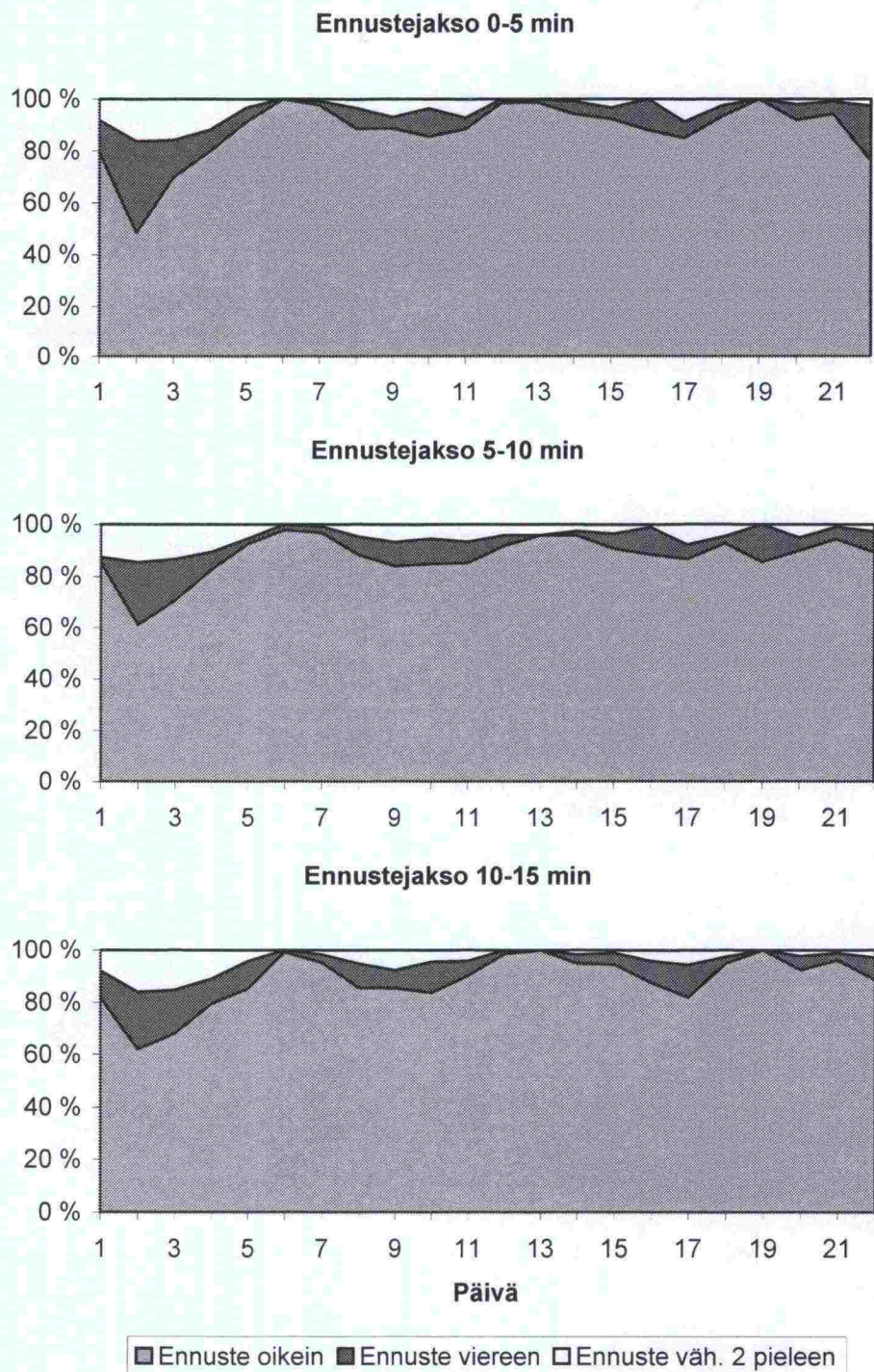
Taulukko 7. Vastejakaumien painotuksen (etäisyyden käänteisluku) vaikutus ennusteeseen. Ennustelinkkinä oli Pakila–Konala, ennustejakso oli 0–5 minuuttia ja aineistona oli testiaineisto.

Ennustettu luokka	3 BMU:ta, ei painotusta					3 BMU:ta, painotus				
	Mitattu luokka									
	1	2	3	4	5	1	2	3	4	5
	1	3244	28	15	0	0	3244	28	15	0
2	2	23	48	0	0	2	22	51	0	0
3	4	7	75	0	0	4	8	72	0	0
4	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0
Ei	21	5	27	0	0	21	5	27	0	0

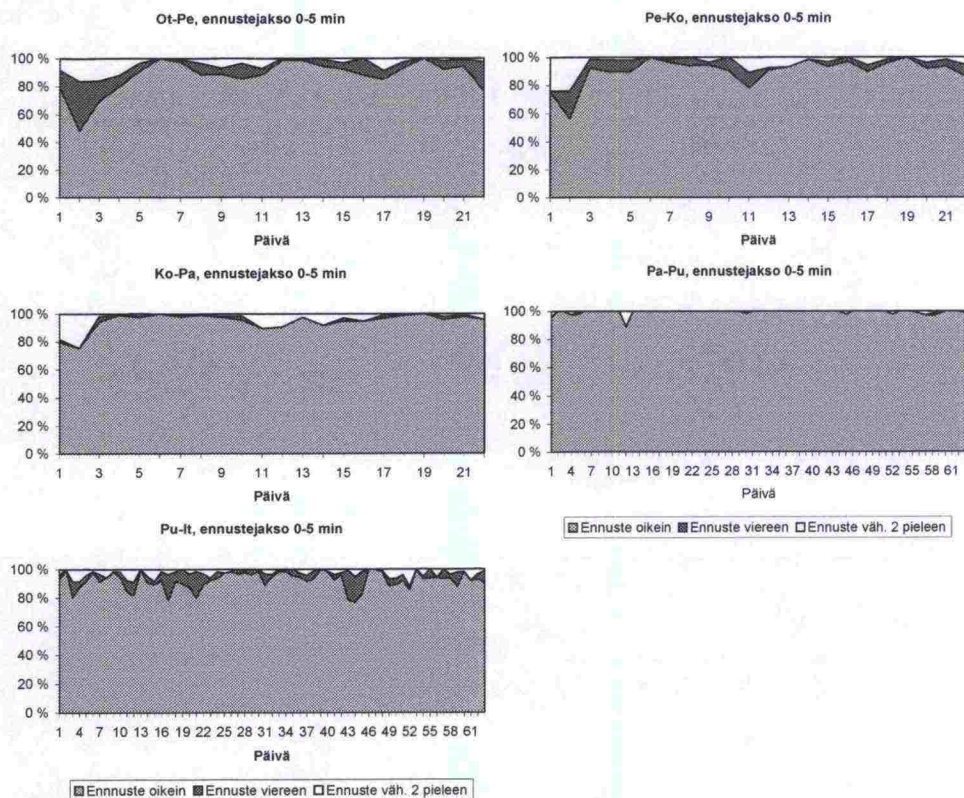
3.4 Online-kokeilun tulokset

Mallin itseoppimisperiaatetta testattiin kesä- ja heinäkuun 2004 aikana. Mallin annettiin toimia ajantasaisesti, ja erilaisten mallin toimintaa kuvaavien tunnuslukujen kehitystä ajan suhteen tarkkailtiin.

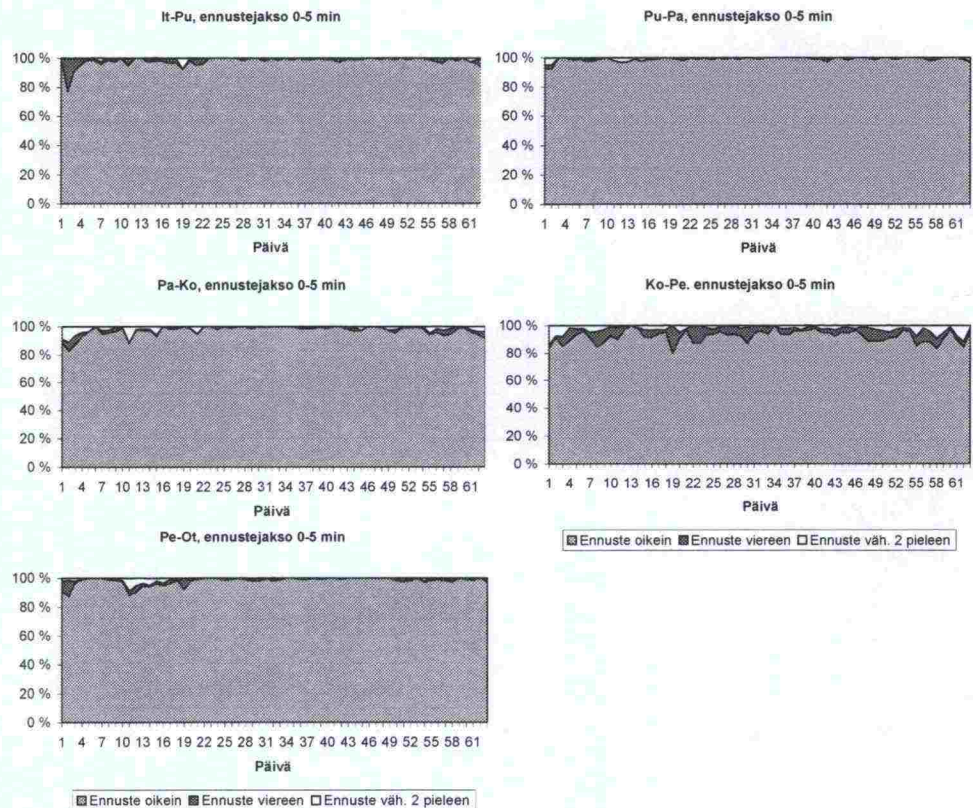
Ennusteiden onnistumista tarkkaillessa havaittiin viikonpäivien vaikutus. Viikonloppuina liikenne oli pääasiassa sujuvaa ja ennusteet onnistuivat paremmin kuin viikolla. Tämä seitsemän päivän sykli on havaittavissa selkeästi. Ennustejaksolla oli vain vähän vaikutusta oikein ennustettujen osuuteen (kuva 12), ja linkkien välinen vaihtelu oli ennustejakson vaikutusta suurempi (kuvat 13 ja 14).



Kuva 12. Online-kokeilun tulokset tiejaksolle Otaniemi-Perkkaa.



Kuva 13. Online-kokeilun tulokset Otaniemi-Itäkeskus-suunnalle. Ennustejakso 0-5 min.



Kuva 14. Online-kokeilun tulokset Itäkeskus– Otaniemi-suunnalle. Ennustejakso 0–5 min.

Online-kokeilun aikana seurattiin myös tekemättä jääneiden ennusteiden osuutta. Niiden ennusteiden osuus, jotka jäivät tekemättä siksi, että klusteri oli tyhjä, oli kokeilun alkuvaiheessa useimmilla osamalleilla noin 20–30 prosenttia. Osuus pieneni kokeilun ajan kaikkien tiejaksojen osamalleilla suunnilleen tahti 0,1 prosenttiyksikköä päivässä. Noin joka toinen ennuste jäi tekemättä jonkin syötesuureen tietojen puuttumisen takia.

3.5 Liikennemäärätiedot lisäsyötteinä

Liikennemäärätiedon tuomaa lisäarvoa testattiin olettamalla tiedon mallille saapumisviive 10 minuutiksi. Liikennemäärätiedot paransivat lisäsyötteinä ennustemalleja useammin kuin huononsivat niitä (taulukko 8). Tulos oli erityisen selkeä Otaniemi–Pakila- ja Pukinmäki–Konala-väleillä. Aina liikennemäärätietokaan ei kuitenkaan paranna ennustetulosta. LAM-pisteverkko on suhteellisen harva (1 mittauspiste/tiejakso) eikä yksittäisen pisteen tilanne aina kuvaa koko tiejakson tilannetta.

Taulukko 8. Pelkkään matka-aikatietoon sekä matka-aika- ja liikennemäärätietoon perustuvien mallien väliset vertailut. Vertailulukuna oikein ennustettujen osuus (%).

		Pelkkä matka-aikatieto			Matka-aika- ja liikennemäärätieto		
		0–5	5–10	10–15	0–5	5–10	10–15
OtPe	Sujuva	77	93	93	80	94	94
	Jonoutunut	69	46	45	72	50	47
	Hidas	91	69	63	89	70	63
	Pysähtelevä	100	80	76	100	80	75
	Seisova	100	0	0	100	0	0
PeKo	Sujuva	87	93	93	90	96	96
	Jonoutunut	83	43	34	87	48	37
	Hidas	86	86	84	85	88	86
	Pysähtelevä	100	94	91	100	96	92
	Seisova	100	100	50	100	100	100
KoPa	Sujuva	90	96	95	92	96	97
	Jonoutunut	93	49	39	96	55	41
	Hidas	94	95	95	94	96	94
	Pysähtelevä	100	68	63	100	71	63
PaPu	Sujuva	92	96	98	88	96	96
	Jonoutunut	87	75	43	94	52	46
	Hidas	94	94	86	96	91	92
	Pysähtelevä	100	100	100	100	100	100
Pult	Sujuva	63	94	96	63	95	95
	Jonoutunut	82	46	38	83	37	33
	Hidas	81	51	46	85	51	46
	Pysähtelevä	100	0	6	100	0	5
ItPu	Sujuva	85	96	96	82	97	98
	Jonoutunut	88	63	61	91	59	57
	Hidas	93	83	74	94	81	76
PuPa	Sujuva	84	93	93	89	96	96
	Jonoutunut	89	53	40	88	51	42
	Hidas	92	92	90	92	94	92
	Pysähtelevä	100	100	0	100	100	0
	Seisova	100	0	0	100	0	0
PaKo	Sujuva	94	96	97	95	98	97
	Jonoutunut	90	50	44	90	56	46
	Hidas	86	90	86	89	90	90
	Pysähtelevä	100	97	88	100	94	96

		Pelkkä matka-aikatieto			Matka-aika- ja liikennemäärätieto		
		0–5	5–10	10–15	0–5	5–10	10–15
KoPe	Sujuva	91	97	97	92	98	97
	Jonoutunut	91	55	38	92	52	49
	Hidas	91	93	90	91	92	90
	Pysähtelevä	100	90	90	100	93	83
PeOt	Sujuva	61	95	96	62	95	95
	Jonoutunut	75	33	28	75	33	29
	Hidas	99	45	38	90	45	42
	Pysähtelevä	100	0	8	100	20	0

Liikennemäärätiedot eivät kuitenkaan saapuneet aina mallille 10 minuutin viiveellä (maksimissaan 10 minuutin viive toteutui noin 9 prosenttia ajasta). Tästä syystä lisäsyötteiden tuomaa hyötyä tarkasteltiin myös pidemmillä viiveillä (15 ja 20 minuuttia). Kun viive kasvoi suuremmaksi kuin 10 minuuttia, liikennemäärätiedosta ei enää ollut hyötyä mallille (taulukko 9).

Taulukko 9. Erimittaisella viiveellä tulevasta liikennemäärätiedosta saatava hyöty. Tulokset ovat linkiltä Pakila–Konala ja ennustejaksonea oli 0–5 minuuttia.

	Oikein ennustettujen osuudet (%) sujuvuusluokittain				
	1	2	3	4	5
Liikennemäärätieto 15 min viiveellä	100	3	78	4	–
Liikennemäärätieto 20 min viiveellä	99	2	77	4	–
Ei liikennemäärätietoa	100	1	79	0	–

3.6 Kelin vaikutus

Kelin vaikutusta ennusteisiin selvitettiin tekemällä ennusteet ilman kelitietoja ja selvittämällä, eroaako mallin suorituskkyky eri tilanteissa. Valtatiellä 3 on tässä käytetyn tiesääaseman paikkeilla – niin kuin Kehä I:lläkin – tien suuresta liikennemäärästä ja hyvästä kunnossapidosta johtuen erittäin vähän huonoa ja etenkin erittäin huonoa keliä.

Tuloksista (liite 1) havaittiin, että mallit olivat keskimäärin suunnilleen yhtä hyviä normaalilla ja erittäin huonolla kelillä, kun tarkastellaan mallien (yksittäinen tiejakso, keskiarvo sen kaikista ennustejaksoista) toimintaa kussakin sujuvuusluokassa erikseen. Sujuvan liikenteen ennustaminen kuitenkin onnistui erittäin huonolla kelillä huonommin kuin normaalilla, muuten päinvastoin. Huonolla kelillä malli ennusti suunnilleen samalla tavalla kuin normaalilla kelillä. Tällöin tulos oli yleensä normaalin ja erittäin huonon kelin välissä.

Vaikka huono ja erittäin huono keli ovatkin harvinaisia, päätettiin malli jakaa osamalleihin kelin perusteella. Kaikilla keleillä sovellettiin tietyn tiejakson tietylle ennustejaksolle samaa itseorganisoituvaa karttaa, mutta kunkin keliluokan havainnoista koottiin omat todennäköisyystaulunsa.

Ajantasaista ennustemallia varten tehtiin kuitenkin päätös, että jos tietyn keli- luokan todennäköisyystaulussa ei ollut havaintoja, ennuste tehtiin yhtä luokkaa paremman kelin todennäköisyystaulun perusteella. Näin pyrittiin turvaamaan ennusteiden syntyminen huonon kelin harvinaisuudesta riippumatta.

3.7 Viikonpäivän vaikutus

Viikonpäivän vaikutusta ennusteisiin selvitettiin vastaavalla tavalla kuin kelin vaikutusta eli tekemällä ennusteet ilman viikonpäivätietoja ja selvittämällä, eroaako mallin suorituskyyky eri tilanteissa.

Viikonloppuina luokan sujuva osuus havainnoista oli erittäin suuri (taulukko 10). Koska viikonloppujen sujuva liikenne ennustettiin keskimäärin paremmin kuin arkisujuvat (liite 1), päätettiin viikonpäivätieto jättää mallista pois.

Taulukko 10. Sujuvan liikenteen osuus (%) eri linkeillä viikonloppujen liikenteestä.

Linkki	Osuus	Linkki	Osuus
OtPe	95,9	ItPu	99,8
PeKo	99,8	PuPa	98,6
KoPa	99,7	PaKo	99,8
PaPu	100,0	KoPe	100,0
PuIt	96,2	PeOt	95,9

3.8 Puuttuva tieto

Ajantasaisesti toimivalle mallille on eduksi, jos se kykenee tekemään ennusteita myös silloin, kun se ei saa kaikkia syötetietoja. Puuttuvien tietojen kanssa toimimiseen on vähintään kolme toimintatapaa:

1. Kullekin osittaisen syötteen yhdistelmälle tehdään oma, vain tiettyihin syötemuuttujiin pohjautuva SOM.
2. Muuttujan arvo pidetään vakiona, kunnes saadaan uusi arvo tai viimeisin mittaustulos on liian vanha.
3. BMU haetaan ilman kyseistä muuttujaa (minimi määritetään vain olemassa olevien muuttujien perusteella).

Mallissa on jo perusasetelmassa kymmenen linkkiväliä, kolme ennustejaksoa ja kolme keliluokkaa, eli 30 karttaa ja 90 todennäköisyystaulukkoa. Vaihtoehto 1 tekee mallin rakenteesta varsin monimutkaisen, sillä sekä karttojen että todennäköisyystaulukoiden määrä moninkertaistuisi.

Vaihtoehto 2 on toimiva silloin, kun liikennetilanne on suhteellisen vakio. Kun ruuhka on kasvamassa tai helpottamassa, vanhentunut arvo lieenee ennusteen kannalta huonompi vaihtoehto kuin kyseisen muuttujan huomiotta jättäminen (vaihtoehto 3). Koska ennusteet ovat erityisen tärkeitä juuri näissä muutostilanteissa, päätettiin osittaisen syötteen haaste ratkaista vaihtoehdon 3 mukaisesti.

Ne vajaiden syötteiden yhdistelmät, jotka hyväksytään malliin, voidaan valita esimerkiksi joko niiden ennusteiden osuuden perusteella, jotka päättyivät osittaisesta syötteestä huolimatta samaan BMU:hun tai samaan ennusteeseen. Näitä osuuksia ruuhkaisessa (ei sujuva) liikenteessä on kuvattu liitteen 2 kuvissa.

Lopulliseen malliin päätettiin valita ne puutteellisten syötteiden kombinaatiot, jotka johtivat ruuhkaisessa liikenteessä vähintään 80 prosenttia ajasta samaan BMU:hun kuin täydellinen syöte. Tällöin ruuhkaisessa liikenteessä ennustetaan yli 90 prosenttisesti sama sujuvuusluokka kuin täydellisellä syötteellä. Tämä on seurausta siitä, että itseorganisoituvan kartan perusperiaatteiden mukaisesti samantyyppiset tilanteet sijoittuvat kartalla lähekkäin, jolloin täydellisen syötteen BMU:n läheisyyteen osuva syöte saattaa johtaa samaan ennusteeseen kuin täydellinen syötekin eri BMU:sta huolimatta.

3.9 Tuotantomallin ohjelmoiminen

Tuotantomalli tehtiin tammi–elokuussa 2004 kootun matka-aika-aineiston pohjalta raportissa kuvatuin periaattein. Mallissa oli kullekin tiejaksolle ja kullekin ennustejaksolle omat itseorganisoituvat kartat (yhteensä 30 kpl) sekä jokaista karttaa kohti kolme todennäköisyystaulukkoa – yksi kullekin keliluokalle (yhteensä 90 taulukkoa).

Mallit toimivat myös osittaisilla syötteillä niissä tilanteissa, jolloin puuttuva syöte ei vaikuttanut voimakkaasti BMU:hun tai ennusteeseen. Sallittuja puuttuvien syötteiden yhdistelmiä oli tieosasta riippuen 0–441.

Sujuvuusennusteen pilottiversion käyttöliittymästä haluttiin yksinkertainen, mutta informatiivinen. Kehä I on esitetty siinä palkein, joiden väri kertoo kulloinkin voimassa olevan ennusteen (liite 3) kullekin ennustejaksolle. Värit (vihreä, sininen, keltainen, oranssi ja punainen) ovat samat kuin Tiehallinnon muussa liikennetiedotuksessa.

Malli ohjelmoitiin päivittämään todennäköisyystaulukoita (itseoppimisperiaatte) raportissa aiemmin kuvattujen periaatteiden mukaisesti. Päivittäminen tehdään kuitenkin ainoastaan silloin, kun syöte on täydellinen. Vaikka puutteellinen syöte siis johtaisikin ennusteeseen, taulukoita ei päivitetä eli itseoppimista ei tapahdu. Vaikka ennuste perustuu tietyissä tilanteissa todellisuutta parempaan keliluokkaan, päivitys tehdään kuitenkin todellisen keliluokan taulukkoon.

Malli kokoaa tapahtumista lokia, johon kirjataan syötteiden lisäksi aika sekä kullekin ennusteajakajaksolle keliluokka, ennuste, toteutunut sujuvuusluokka, BMU ja ennusteen pohjalla ollut todennäköisyysjakauma.

4 TULOSTEN TARKASTELU JA PÄÄTELMÄT

Tutkimuksen tavoitteena oli tehdä itseoppiva lyhyen aikavälin sujuvuusennustemalli. Tarkoituksena oli tehdä malli, joka ennustaa tiejaksoittain seuraavan 15 minuutin kuluessa lähdössä olevien ajoneuvojen sujuvuusluokan viisiportaisen luokittelun mukaisesti viiden minuutin jaksoissa Kehä I:lle. Toisaalta tarkoituksena oli kehittää malli sellaiseksi, että se oppii itse kohtaamistaan liikennetilanteista ja sopeuttaa ennusteitaan niiden perusteella ilman, että kaikkea mitattua aineistoa tarvitsee tallettaa.

Tutkimuksen tavoitteet saavutettiin. Työn tuloksena saatiin kehitettyä malli, joka kykeni ennustamaan tiejaksojen sujuvuuden. Rakenteensa ansiosta malli kykeni oppimaan kohtaamistaan liikennetilanteista ilman, että kaikkea liikennetietoa piti tallettaa tietokantoihin. Tämän mahdollisti havaintojen luokittelu ja kunkin luokan vastetodennäköisyystaulukoiden päivittäminen.

Mallia tehdessä kävi selväksi, ettei pelkkä mediaani kyennyt suodattamaan kaikkia poikkeavien matka-aikahavaintojen aiheuttamia vääristymiä. Yksinkertainen menetelmä osoittautui tehokkaaksi aineiston ajantasaisesti tehtävässä esikäsittelyssä: mediaaniaineisto suodatettiin havaintojen lukumäärän ja sen perusteella, kuinka paljon saatu mediaani suhteellisesti poikkesi edellisestä hyväksytystä arvosta.

Online-kokeilun mukaan niiden ennusteiden osuus, jotka jäivät tekemättä siksi, että klusteri oli tyhjä, pieneni itseoppimisperiaatteen ansiosta ajan myötä. Pieneneminen oli kaikkien linkkien osamalleilla suunnilleen 0,1 prosenttiyksikköä päivässä.

Malli olisi monissa tilanteissa ennustanut sujuvuutta paremmin, jos sillä olisi ollut käytettävissään matka-aikojen lisäksi tieto liikennemäärästä. Nyt liikennemäärätieto oli käytettävissä, mutta se tuli niin pitkällä viiveellä (usein jopa 20 minuuttia), että tämä viive kumosi tiedon tuomat edut. Matka-aika on luonteeltaan sellainen, että se pysyy pitkään suhteellisen samalla tasolla, kunnes liikennemäärän kasvaessa tietyn kriittisen rajan yli kasvaa nopeasti. Tämän tilanteen ennakoiminen pelkän matka-aika-aineiston avulla on hankalaa, ja ajantasainen tieto liikennemäärästä tukisi ennustetyötä huomattavasti.

Malli kykenee toimimaan tietyillä osittaisilla syötteillä jo nykyisellään. Jäljelle jää kuitenkin varsin paljon tilanteita, jolloin oikeaa karttayksikköä (BMU) ei kyetä määrittämään riittävän luotettavasti eikä ennustetta näin ollen kannata antaa. Jos malliin olisi kehitetty kullekin puutteelliselle syötetyypille oma itseorganisoituva kartta, joka perustuisi vain kulloinkin käytettävissä oleviin suureisiin, mallin ennustekyky olisi mahdollisesti ollut parempi jossain näistä tilanteista. Tämän tyyppinen osamallitus tekee kuitenkin ennustemallin rakenteesta monimutkaisemman ja se vaatisi lisäkehitystyötä, jotta malli kykenisi tästä huolimatta toimimaan kuormittamatta tietokonetta kohtuuttomasti.

Toinen kehitysidea liittyy mallin käyttöliittymään. Nyt ennusteet esitettiin peräkkäin värikoodein ja todennäköisyyksin peräkkäin asetettujen palkkien avulla. Jos ennusteita halutaan esittää esimerkiksi Internetissä, esitystavan helppolukuisuutta ja näytävyyttä voisi lisätä karttapohjan avulla.

Nyt Kehä I:lle kehitetyn mallin toimintaperiaate on sellainen, että se on sellaisenaan siirrettävissä helposti myös muualle. Kussakin kohteessa olemassa oleva liikenteen seurantajärjestelmä sanelee syötesuurevaihtoehdot. Alue, jolta yksittäisen osamallin syötteet kootaan, on riippuvainen paikasta ja sille tyypillisistä liikenteen ominaisuuksista. Syötesuureita voidaan ensin rajata karkeammin asiantuntija-arvauksen pohjalta, ja lopulliset syötteet voi selvittää raportissa esitetyllä tavalla.

KIRJALLISUUSLÄHTEET

ELORANTA T (1999). Rekisterikilpien tunnistukseen perustuva liikenteen automaattinen matkanopeuden seuranta. Tielaitoksen selvityksiä 46/1999. Uudenmaan tiepiiri, Tielaitos, Helsinki. 149 s.

INNAMAA S (2004). Matka-ajan ajantasainen ennustaminen, Pilottikokeilu. Tutkimusraportti RTE474/04, VTT Rakennus- ja yhdyskuntatekniikka, Espoo. 26 s.

KILJUNEN M, SUMMALA H (1996). Ruuhkaisuuden kokeminen ja liikennetilanne-tiedottaminen. Tienkäyttäjätutkimus kaksikaistaisilla teillä. *Tielaitoksen selvityksiä* 25/1996, Tiehallinto, Helsinki. 77 + 5 s.

KOHONEN T (2001). Self-organizing maps. Third edition. Springer series in information sciences. 501 s.

LUOMA S (1998). Tieliikenteen sujuvuus ja sen mittaaminen. Diplomityö, Teknillinen korkeakoulu, Liikennetekniikka. 122 s.

OHBA Y, UENO H, KUWAHARA M (2000). Travel time prediction method for expressway using toll collection system data. Proceedings of the 7th World Congress on Intelligent Systems. 8 s.

RAITIO J (1998). Sään ja kelin tunnistaminen muuttuvien nopeusrajoitusten ohjaamiseksi Kotka–Hamina-moottoritieellä. Diplomityö, Teknillinen korkeakoulu, Espoo. 78 s.

VESANTO J, HIMBERG J, ALHONIEMI E, PARHANKANGAS J (2000). SOM Toolbox for Matlab 5. Raportti A57, SOM Toolbox Team, Teknillinen korkeakoulu, Espoo. 59 s.

LIITTEET

- LIITE 1 Kelin ja viikonpäivän vaikutus ennusteisiin
- LIITE 2 Puutteellinen syöte ja ennusteen onnistuminen
- LIITE 3 Pilottiversion käyttöliittymä
- LIITE 4 Käyttöohje

KELIN JA VIIKONPÄIVÄN VAIKUTUS ENNUSTEISIIN

Taulukko 1. Oikeaan osuneiden ennusteiden osuudet (%) keli- ja sujuvuusluokittain.

Linkki	Keli	Sujuvuus	Ennustejakso		
			0–5 min	5–10 min	10–15 min
OtPe	normaali	sujuva	80	94	95
	huono		79	94	94
	erittäin huono		78	96	96
	normaali	jonoutunut	72	51	47
	huono		72	46	46
	erittäin huono		77	44	31
	normaali	hidas	90	72	66
	huono		90	65	58
	erittäin huono		85	69	59
PeKo	normaali	sujuva	90	96	97
	huono		88	95	95
	erittäin huono		87	94	93
	normaali	jonoutunut	88	48	37
	huono		86	48	40
	erittäin huono		94	38	29
	normaali	hidas	85	88	87
	huono		86	90	85
	erittäin huono		96	93	74
KoPa	normaali	sujuva	93	97	97
	huono		89	95	97
	erittäin huono		79	94	96
	normaali	jonoutunut	96	56	41
	huono		94	52	41
	erittäin huono		100	56	48
	normaali	hidas	94	96	95
	huono		92	95	93
	erittäin huono		60	100	57

Linkki	Keli	Sujuvuus	Ennustejakso		
			0–5 min	5–10 min	10–15 min
PaPu	normaali	sujuva	89	96	96
	huono		82	94	94
	erittäin huono		57	91	83
	normaali	jonoutunut	93	74	50
	huono		100	17	29
	erittäin huono		100	100	
	normaali	hidas	96	93	91
	huono		90	100	100
	erittäin huono				
Pult	normaali	sujuva	63	95	95
	huono		64	95	95
	erittäin huono		54	93	96
	normaali	jonoutunut	83	39	33
	huono		81	17	27
	erittäin huono		91	38	50
	normaali	hidas	86	55	51
	huono		82	17	7
	erittäin huono		67	67	75
ItPu	normaali	sujuva	83	97	98
	huono		76	97	98
	erittäin huono		57	95	96
	normaali	jonoutunut	87	48	49
	huono		93	63	61
	erittäin huono		93	87	66
	normaali	hidas	90	64	53
	huono		95	84	79
	erittäin huono		100	93	100
PuPa	normaali	sujuva	89	96	97
	huono		85	95	96
	erittäin huono		82	92	90
	normaali	jonoutunut	89	51	39
	huono		88	52	45
	erittäin huono		83	49	60
	normaali	hidas	92	94	92
	huono		91	93	90
	erittäin huono		96	94	94

LIITTEET

Linkki	Keli	Sujuvuus	Ennustejakso		
			0–5 min	5–10 min	10–15 min
PaKo	normaali	sujuva	96	98	98
	huono		92	97	97
	erittäin huono		84	94	93
	normaali	jonoutunut	88	53	43
	huono		93	58	48
	erittäin huono		100	76	68
	normaali	hidas	89	90	90
	huono		91	89	90
	erittäin huono		84	90	94
KoPe	normaali	sujuva	93	98	98
	huono		88	96	96
	erittäin huono		78	95	93
	normaali	jonoutunut	89	56	50
	huono		93	49	48
	erittäin huono		100	37	44
	normaali	hidas	88	88	85
	huono		94	93	94
	erittäin huono		96	98	94
PeOt	normaali	sujuva	64	95	96
	huono		55	93	94
	erittäin huono		59	92	89
	normaali	jonoutunut	75	27	23
	huono		75	41	40
	erittäin huono		85	73	68
	normaali	hidas	98	39	35
	huono		100	44	43
	erittäin huono		100	88	91

Taulukko 2. Oikeaan osuneiden ennusteiden osuudet (%) viikonpäivittäin ja sujuvuusluokittain.

Linkki	Sujuvuus	Päivä	Ennustejakso		
			0–5 min	5–10 min	10–15 min
OtPe	sujuva	ma	60	94	95
		ti	62	94	94
		ke	59	94	95
		to	60	94	95
		pe	61	94	95
		la	68	98	96
		su	69	97	97
	jonoutunut	ma	79	39	34
		ti	76	36	34
		ke	79	38	34
		to	76	31	24
		pe	80	31	28
		la	53	14	9
		su	54	14	20
	hidas	ma	100	54	49
		ti	98	53	56
		ke	100	63	56
		to	96	39	45
		pe	100	4	13
		la	100	13	0
		su	94	13	0
PeKo	sujuva	ma	87	95	96
		ti	85	94	95
		ke	86	94	95
		to	87	95	95
		pe	87	94	94
		la	98	99	99
		su	99	100	100
	jonoutunut	ma	88	47	44
		ti	84	57	37
		ke	89	52	35
		to	89	47	41
		pe	88	38	30
		la	50	0	0
		su	0	0	

LIITTEET

Linkki	Sujuvuus	Päivä	Ennustejakso		
			0–5 min	5–10 min	10–15 min
PeKo	hidas	ma	87	84	82
		ti	84	86	86
		ke	87	88	87
		to	85	92	90
		pe	82	92	89
		la	0	57	0
		su			
KoPa	sujuva	ma	90	95	97
		ti	91	96	97
		ke	90	96	97
		to	91	95	97
		pe	90	95	96
		la	94	98	99
		su	94	98	99
	jonoutunut	ma	100	51	37
		ti	97	62	54
		ke	97	68	44
		to	94	55	33
		pe	92	26	30
		la	89	20	25
		su	50	0	0
	hidas	ma	94	97	93
		ti	91	95	94
		ke	93	96	93
		to	96	97	96
		pe	94	97	96
		la	50	0	0
		su	0	0	0
PaPu	sujuva	ma	85	95	95
		ti	86	95	96
		ke	84	94	95
		to	85	95	96
		pe	88	95	95
		la	97	99	99
		su	92	96	97

Linkki	Sujuvuus	Päivä	Ennustejakso		
			0–5 min	5–10 min	10–15 min
PaPu	jonoutunut	ma	100	50	0
		ti	100	0	50
		ke	100	55	33
		to	100	78	56
		pe	50	100	100
		la			
		su			
	hidas	ma	89	94	88
		ti	85	92	89
		ke	100	82	89
		to	100	89	88
		pe	100	100	97
		la			
		su			
Pult	sujuva	ma	60	94	94
		ti	63	96	96
		ke	61	93	95
		to	62	95	95
		pe	63	95	95
		la	66	96	94
		su	70	96	95
	jonoutunut	ma	89	34	35
		ti	78	30	21
		ke	83	47	43
		to	80	46	34
		pe	88	27	31
		la	77	24	15
		su	63	21	25
	hidas	ma	82	16	19
		ti	70	25	22
		ke	86	55	50
		to	90	52	57
		pe	89	74	59
		la	67	0	0
		su	64	8	0

Linkki	Sujuvuus	Päivä	Ennustejakso		
			0–5 min	5–10 min	10–15 min
ItPu	sujuva	ma	76	96	97
		ti	79	96	98
		ke	77	96	96
		to	78	96	97
		pe	79	97	98
		la	93	99	100
		su	92	98	100
	jonoutunut	ma	93	68	63
		ti	90	56	60
		ke	91	65	63
		to	89	55	56
		pe	91	53	45
		la	71	67	14
		su	75	0	0
	hidas	ma	98	78	81
		ti	98	83	81
		ke	94	82	77
		to	88	96	77
		pe	93	80	75
		la	75	0	0
		su	100	0	0
PuPa	sujuva	ma	86	95	96
		ti	87	95	96
		ke	87	95	96
		to	85	95	95
		pe	85	94	94
		la	95	99	99
		su	95	99	99
	jonoutunut	ma	90	50	42
		ti	90	45	34
		ke	89	55	51
		to	87	58	41
		pe	88	51	42
		la	78	10	3
		su	86	55	52

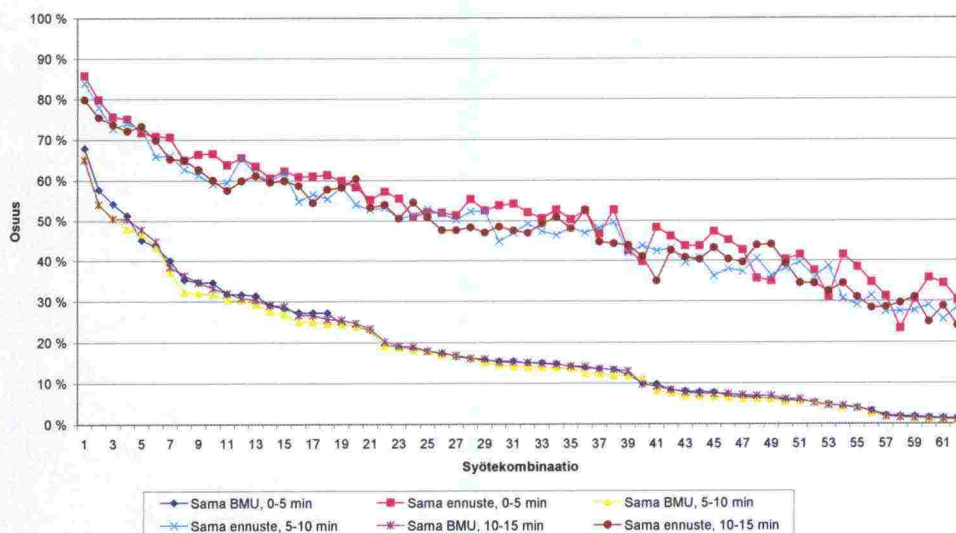
Linkki	Sujuvuus	Päivä	Ennustejakso		
			0–5 min	5–10 min	10–15 min
PuPa	hidas	ma	94	95	93
		ti	92	94	90
		ke	94	94	93
		to	93	93	92
		pe	89	93	93
		la	0	0	0
		su	63	56	50
PaKo	sujuva	ma	94	98	97
		ti	94	97	97
		ke	94	97	97
		to	94	97	97
		pe	94	97	96
		la	98	100	100
		su	98	100	99
	jonoutunut	ma	93	64	46
		ti	88	35	34
		ke	94	61	53
		to	87	47	42
		pe	90	67	52
		la	75	17	25
		su	80	33	17
	hidas	ma	89	88	89
		ti	92	92	92
		ke	88	89	89
		to	91	92	91
		pe	86	88	89
		la	0	0	0
		su	67	0	0
KoPe	sujuva	ma	90	98	97
		ti	89	97	96
		ke	89	97	96
		to	91	97	97
		pe	92	98	97
		la	98	100	100
		su	98	100	100

Linkki	Sujuvuus	Päivä	Ennustejakso		
			0–5 min	5–10 min	10–15 min
KoPe	jonoutunut	ma	90	57	53
		ti	94	53	39
		ke	95	46	42
		to	94	62	74
		pe	91	39	35
		la	0	0	
		su	33	0	50
	hidas	ma	92	90	90
		ti	92	94	89
		ke	93	93	91
		to	79	89	89
		pe	90	87	89
		la			
		su			
PeOt	sujuva	ma	60	94	95
		ti	62	94	94
		ke	59	94	95
		to	60	94	95
		pe	61	94	95
		la	68	98	96
		su	69	97	97
	jonoutunut	ma	79	39	34
		ti	76	36	34
		ke	79	38	34
		to	76	31	24
		pe	80	31	28
		la	53	14	9
		su	54	14	20
	hidas	ma	100	54	49
		ti	98	53	56
		ke	100	63	56
		to	96	39	45
		pe	100	4	13
		la	100	13	0
		su	94	13	0

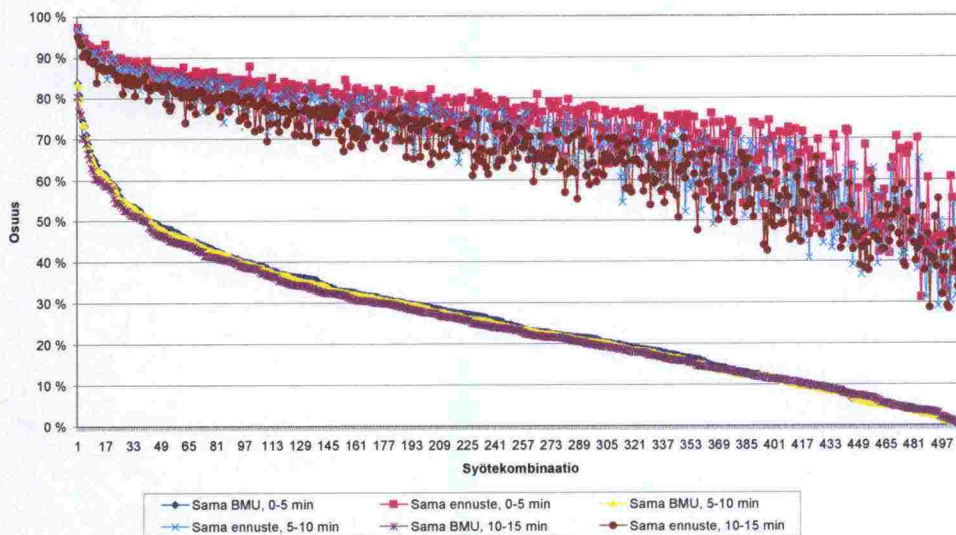
PUUTTEELLINEN SYÖTE JA ENNUSTEEN ONNISTUMI- NEN

Kuvaajat perustuvat ei-sujuvaan liikenteeseen. Kuvissa on X-akselilla erilaiset syöteyhdistelmien (puuttuvien syötteiden määrän ja sijainnin kaikki variaatiot) järjestysnumero. Y-akselilla on saman ennusteen ja saman BMU:n osuudet täydelliseen syötteeseen verrattuna ruuhkaisen liikenteen havainnoissa.

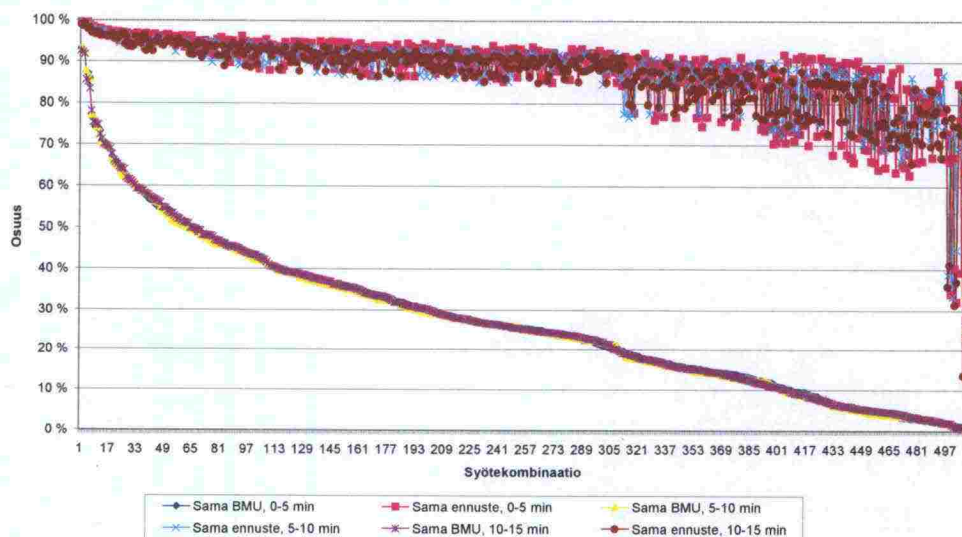
Otaniemi-Perkkaa



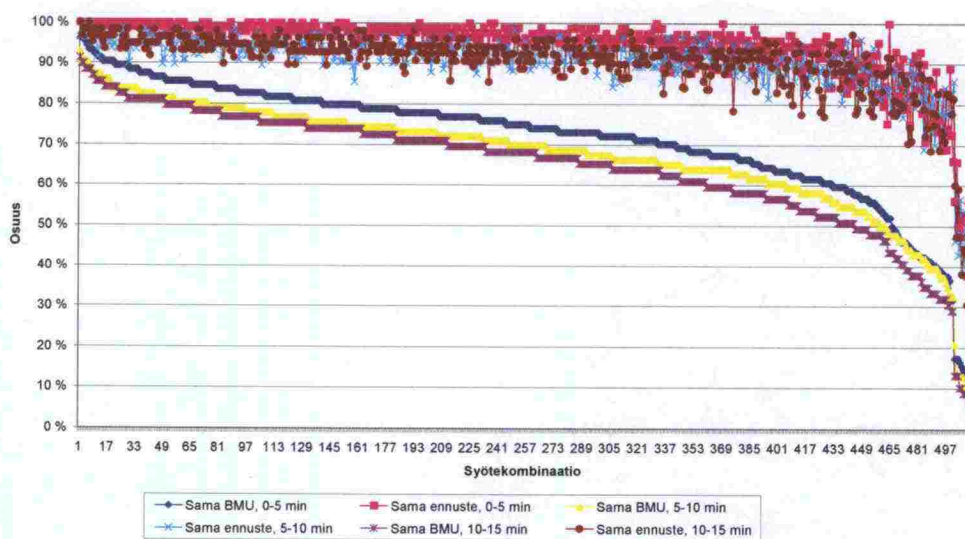
Perkkaa-Konala



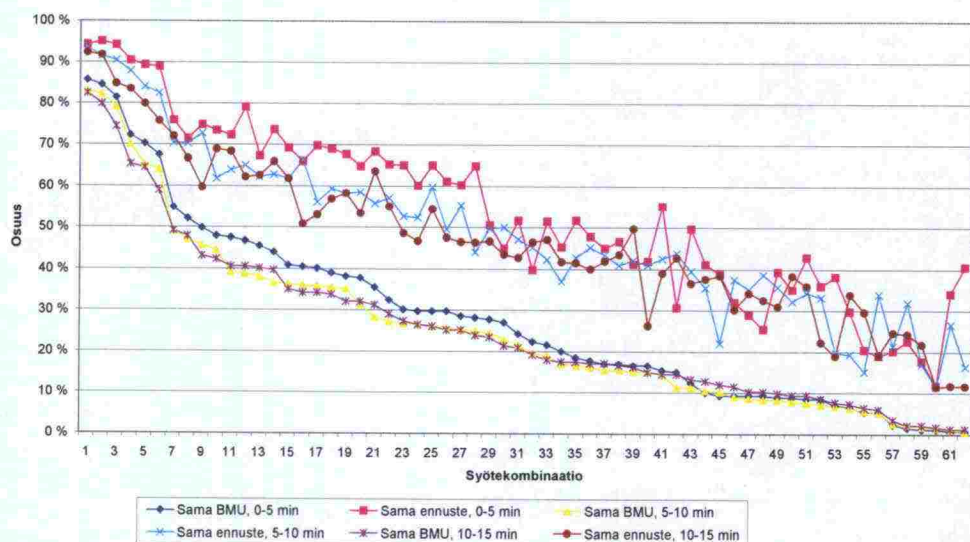
Konala-Pakila



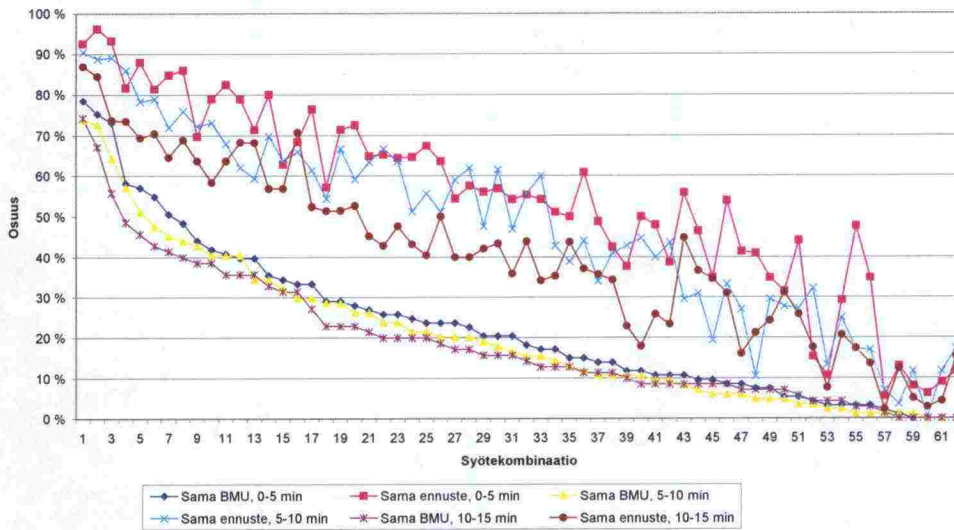
Pakila-Pukinmäki



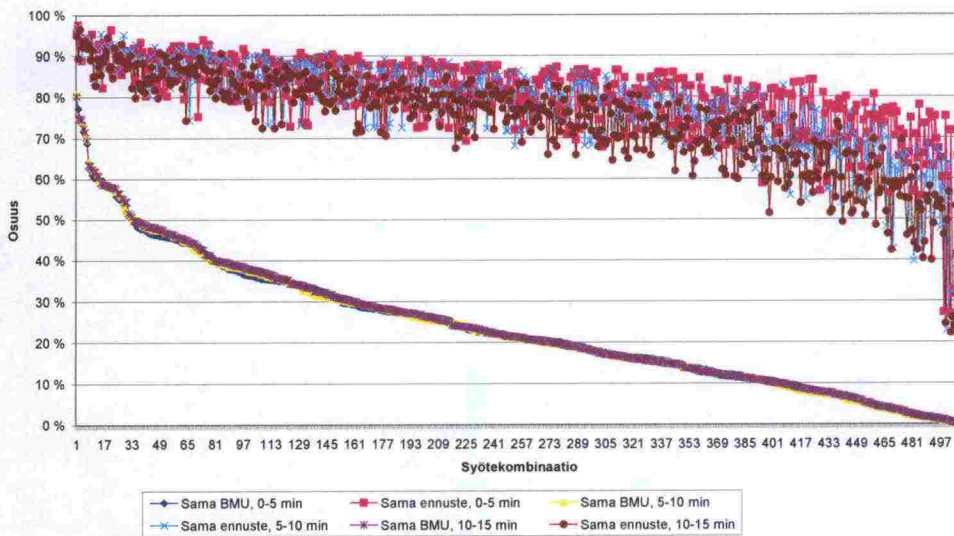
Pukinmäki-Iitäkeskus



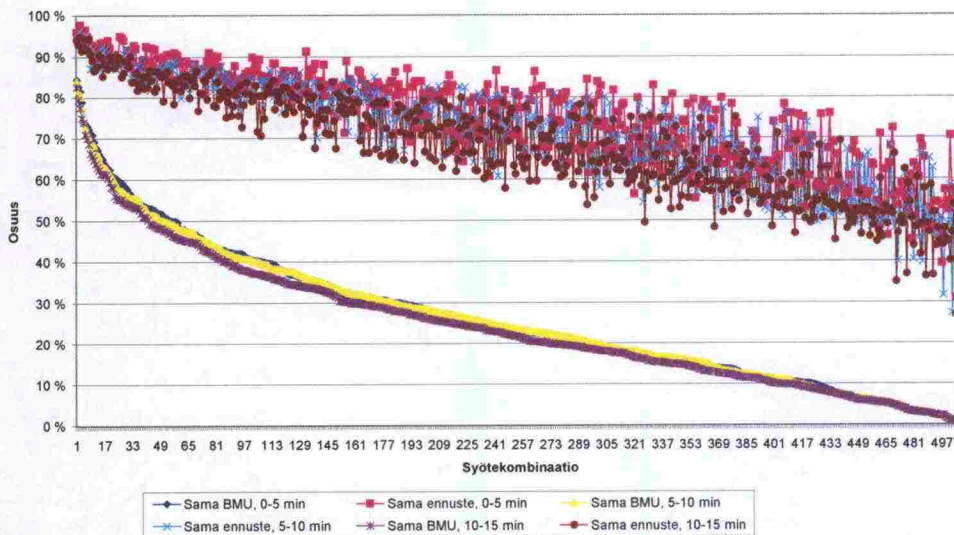
Itäkeskus-Pukinmäki



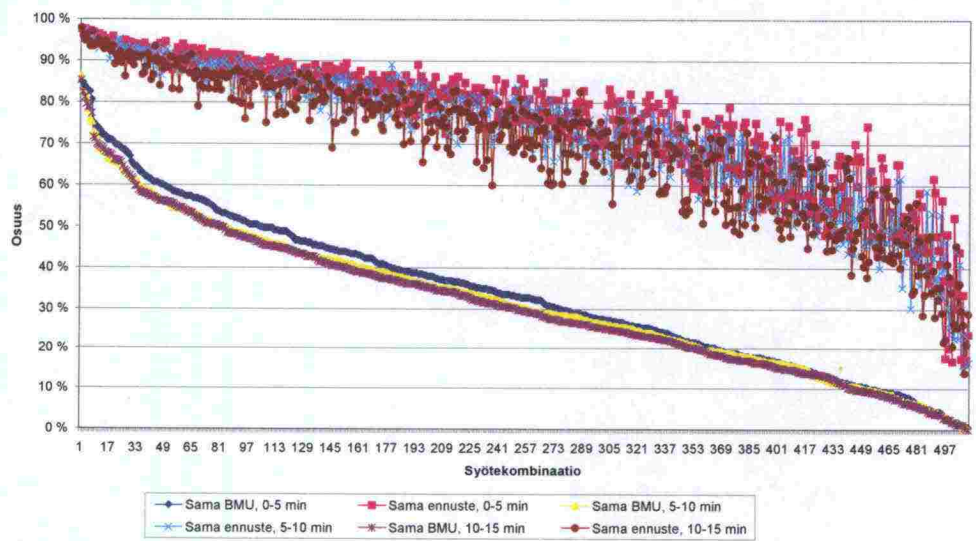
Pukinmäki-Pakila



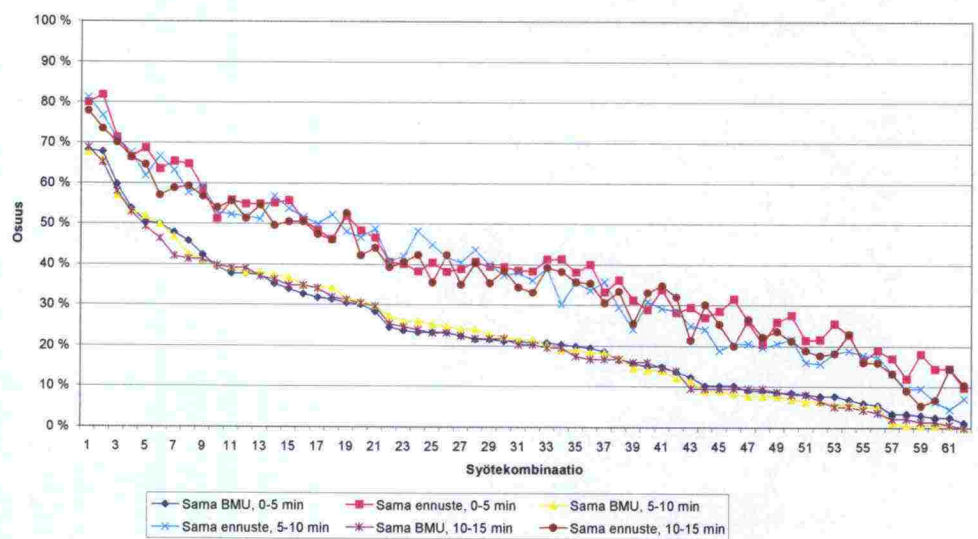
Pakila-Konala



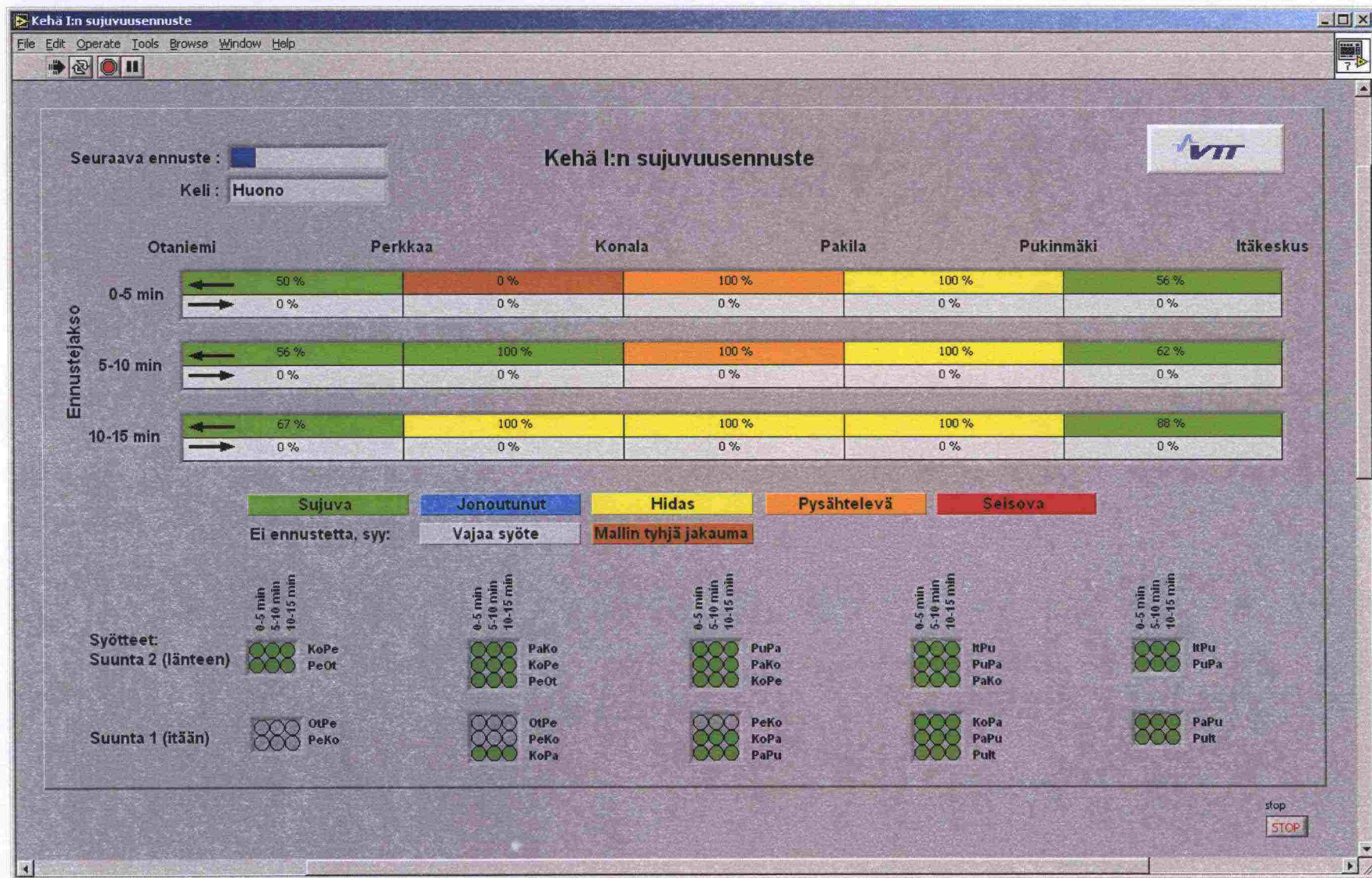
Konala-Perkkaa



Perkkaa-Otaniemi



PILOTTIVERSION KÄYTTÖLIITTYMÄ



KÄYTTÖOHJE

Miten sujuvuusluokat on määritelty?

Sujuvuusluokat on määritelty matkanopeuden ja vapaan nopeuden välisen suhteen perusteella.

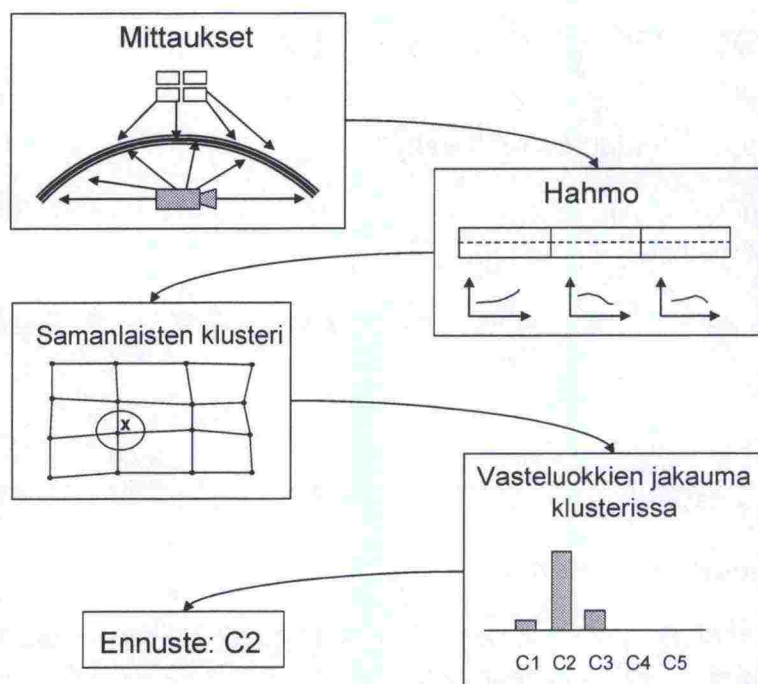
Sujuvuusluokka	Matka-nopeuden ja vapaan nopeuden osamäärä (%)
Sujuva	> 90
Jonoutunut	75–90
Hidas	25–75
Pysähtelevä	10–25
Seisova	< 10

Miten ennuste eroaa ALKin sujuvuustiedosta?

ALKissa esitetyt sujuvuudet perustuvat LAM-pisteiden nopeustietoihin ja tämän mallin antamat ennusteet perustuvat matka-aikojen mittausjärjestelmän tuottamiin matkanopeustietoihin. ALKissa esitetty tieto kuvaa siis yksittäisen pisteen liikennetilannetta ja sujuvuusennuste koko tiejakson keskimääräistä tilannetta. Jotta tiejakson sujuvuus laskee alempaan luokkaan, täytyy koko tiejakson matka-ajan laskea raja-arvon alle. Vaikka yksittäisessä tiejakson pisteessä nopeus olisikin hyvin alhainen, loppujakson sujuvampi liikenne saattaa kompensoida sen ja koko tiejakson sujuvuusluokka olla tätä yksittäistä pistettä korkeampi. Sama voi tapahtua myös toisinpäin.

Miten malli toimii?

1. Kootaan ennustelinkin sekä sitä edeltävän ja seuraavan linkin matka-aikamediaaneista aikasarjat (3 viimeistä 5 minuutin mediaania eli viimeisen 15 minuutin liikennetilanne linkillä). Nämä muodostavat ns. hahmovektorin.
2. Etsitään itseorganisoiuvalta kartalta (SOM-kartta, eräs neuroverkko) klusteri, joka koostuu samantapaisista havainnoista kuin nyt mitattu hahmovektori.
3. Tutkitaan, mikä on tyypillisin seuraus (vaste) tässä klusterissa tällä kelillä. Tämä vasteluokka muodostaa ennusteen. Ennusteena annetaan ennustelinkille lähdössä olevien ajoneuvojen sujuvuusluokat seuraavalle 15 minuutille 5 minuutin jaksoissa.



Malli on ns. itseoppiva eli se oppii kohtaamistaan liikennetilanteista. Itseoppiminen tapahtuu siten, että kun ns. oikea vastaus saadaan mitattua, malli päivittää vasteluokkien jakaumaa.

Miten ennuste luetaan näytöltä?

Ennusteet on esitetty tiejaksoittain (5/suunta), suunnittain ja ennustejaksoittain (0–5, 5–10 ja 10–15 minuuttia ennusteen tekohetkestä eteenpäin). Ennusteluokat on esitetty värikoodilla (vihreä = sujuva, sininen = jonoutunut, keltainen = hidas, oranssi = pysähtelevä ja punainen = seisova).

Ennusteluokan lisäksi palkista näkee ennusteen luotettavuuden. Yleisimpään vasteluokkaan (ennustettu luokka) kuuluvien havaintojen osuus kaikista kyseisen karttayksikön läheisyydessä olevista havainnoista on esitetty kutakin ennustejaksoa kuvaavassa palkissa.

Ne tilanteet, jolloin ennustetta ei voida tehdä on jaettu kahteen osaan: tilanteet, joissa syöte ei ole ollut riittävän kattava, on esitetty harmaalla ja ne tilanteet, joissa ennuste on jäänyt tekemättä tyhjän todennäköisyysjakauman takia, on esitetty ruskealla. Ruskean osuuden pitäisi ajan mittaan vähetä.

Kunkin osamallin syötteet ja se, onko kustakin muuttujasta tuoretta mittaus-tietoa, on esitetty palkkien alla liikennevaloin (vihreä = tuore tieto, harmaa = ei tuoretta tietoa). Palkkien yläpuolella taas kerrotaan ohjelman tulkinta keli-luokasta sekä aikapalkki, josta voi arvioida ajan seuraavaan ennusteen päi-vityshetkeen (ennusteen päivitys 5 min välein).

Aika seuraavaan ennusteen
päivitykseen
(päivitysväli 5 min)

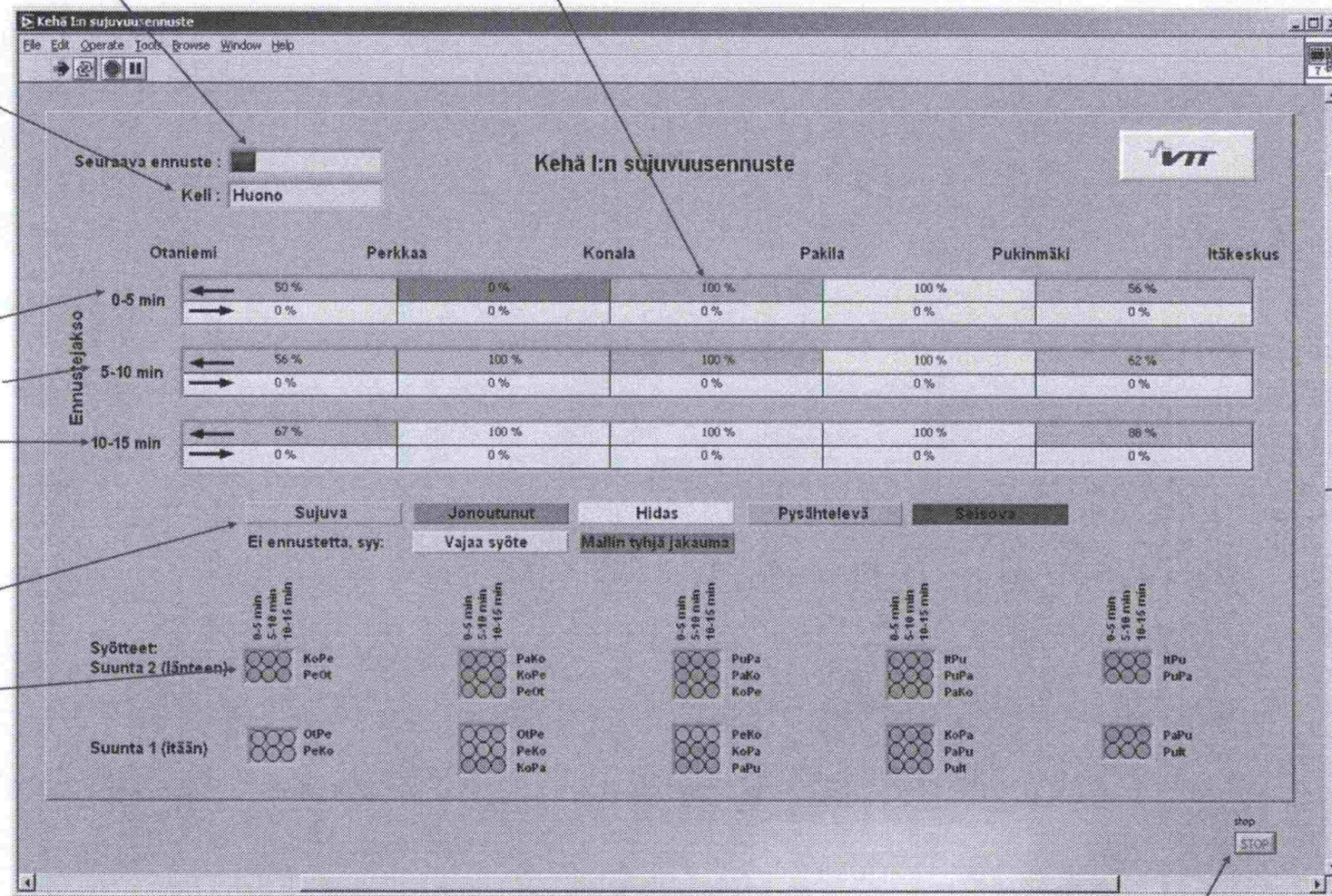
Kuinka suuri osa samantapaisista
tilanteista on johtanut ennustettuun
sujuvuusluokkaan

Keliluokka

Sujuvuusennuste
niille ajoneuvoille, jotka
ovat lähdössä tiejaksolle
seuraavan 5 minuutin
aikana,
5-10 minuutin kuluttua
ja 10-15 minuutin
kuluttua

Värikoodien selitykset

Osamallin syötet
ja se, onko
syötesuurelta saatu
tuore mittaustulos
(vihreä = tuore arvo,
harmaa = mittaustieto
puuttuu)



Ohjelman pysäytys

ISSN 1457-9871
ISBN 951-803-400-1
TIEH 3200910